DOCUMENT RESUME

ED 108 768 95 PS 008 091

AUTHOR Cline, Marvin G.; And Others
TITLE Education as Experimentation: Evaluation of the Follow Through Planned Variation Model. Volume 2A: Two-Year Effects of Follow Through. Final Report.
INSTITUTION Abt Associates, Inc. Cambridge, Mass.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Office of Planning, Budgeting, and Evaluation.
REPORT NO AAI-74-111
PUB DATE 1 May 75
CONTRACT OEC-0-72-5221
NOTE 378p.; For Volumes 1A and 1B of this report, see ED 094 890 and ED 094 891; for Volume 2B, see PS 008 092

EDRS PRICE MF-$0.76 HC-$19.67 PLUS POSTAGE
DESCRIPTORS Comparative Statistics; Compensatory Education Programs; Environmental Influences; Grade 1; Intervention; Kindergarten; *Longitudinal Studies; Measurement Techniques; *Models; *Primary Education; *Program Evaluation; Reliability; Sampling; *Statistical Analysis; Summative Evaluation
IDENTIFIERS *Project Follow Through

ABSTRACT

This second annual report on the National Evaluation of Follow Through (a continuation of the first report) focuses on the accumulated effects of two years of Follow Through experiences on children who remained with the same sponsors over that time period. Comparisons between Follow Through and Non-Follow Through children at the end of first grade are described. In addition to previously used child and school level analyses, site level analyses are included. Chapter 1 presents a brief overview of the report. Chapter 2 describes sample development, including definition of the current sample, comparison with last year's, and discussion of the problems of sample bias, generalizability and sponsor-to-sponsor comparisons. Chapters 3 and 4 deal with the variables used in the analyses and includes a description of covariance selection and the final covariate set. Chapter 5 presents a description of the samples and kindergarten one-year effects for each of the ten program models. Chapter 6 presents site descriptions and effects and overall two-year effects. Chapter 7 presents conclusions of the evaluation. Appendices, which make up approximately one-third of the document, contain data tables and other supportive materials. (ED)

ED108768

Education as Experimentation: A Planned Variation Model

BLACKBOARD ERASER
PHONICS WORKBOOK
Book B

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: PS | TM
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

PS 008091

Abt Associates Inc.

Volume IIA : Two Year Effects of Follow Through

ABT ASSOCIATES INC.
55 WHEELER STREET, CAMBRIDGE, MASSACHUSETTS 02138
TELEPHONE • AREA 617-492-7100
TELEX 710-320-6367

Report No. AAI-74-111

Contract No. OEC-0-72-5221

FINAL REPORT

EDUCATION AS EXPERIMENTATION: EVALUATION OF THE FOLLOW THROUGH PLANNED VARIATION MODEL

VOLUME IIA: TWO-YEAR EFFECTS OF FOLLOW THROUGH

May 1, 1975

Principal Authors:

Marvin G. Cline
Project Director

N. Ames, R. Anderson, M. Conte, T. Ferb,
C. Hall, M. Joshi, J. Larson, D. Park,
E. Proper, L. Stebbins, C. Stern

Prepared for:

THE U.S. OFFICE OF EDUCATION

Office of Planning, Budgeting, and Evaluation
Department of Health, Education, and Welfare
Washington, D.C. 20202

Contract Manager	Quality Control Reviewer	Management Reviewer

EDUCATION AS EXPERIMENTATION:

EVALUATION OF THE FOLLOW THROUGH PLANNED VARIATION MODEL

VOLUME IIA

TWO-YEAR EFFECTS OF FOLLOW THROUGH

CONTENTS

CHAPTER I, OVERVIEW I-1
  1.0 Introduction I-1
  1.1 General Analytic Approach I-2

CHAPTER II, SAMPLE DEVELOPMENT II-1
  2.0 Introduction II-1
  2.1 Sample Selection Criteria II-2
    2.1.1 1974 Analytic Samples II-2
    2.1.2 Reliability Samples II-8
  2.2 Representativeness of 1974 Analytic Sample II-8
    2.2.1 Comparison of Traced Longitudinal Random Sample with Random Sample of Fall 1971 Tested Population II-11
    2.2.2 Comparison of 1974 Analytic Sample with Traced Longitudinal Sample II-14
    2.2.3 Comparison of 1974 Analytic Sample with Fall 1971 Tested Population and 1973 Analytic Sample II-22
  2.3 Discussion II-38

CHAPTER III, VARIABLES III-1
  3.0 Introduction III-1
  3.1 Covariates III-1
  3.2 Outcome Measures III-2
    3.2.1 Tests Administered by SRI III-2
    3.2.2 Domains Examined in the Achievement Batteries III-2

CHAPTER IV, COVARIANCE SELECTION FOR THE SPONSOR WITHIN FT/NFT ANALYSES IV-1
  4.0 Introduction IV-1
  4.1 Initial Covariate Set IV-1
  4.2 Constraints Placed Upon Final Covariate Set IV-5
    4.2.1 Summary of Excluded Covariates IV-8
  4.3 Final Covariate Set IV-11

CHAPTER V, KINDERGARTEN STUDY: A REPLICATION V-1
  5.0 Introduction V-1
  5.1 Far West, Responsive Educational Program V-6
    5.1.1 Sample Description V-6
    5.1.2 Kindergarten 1-Year Effects V-6

5.2 Arizona, Tucson Early Educational Model V-8

  5.2.1 Sample Description V-8

  5.2.2 Kindergerten 1-Year Effects V-9

5.3 Bank Street College Approach V-11

  5.3.1 Sample Description V-11

  5.3.2 Kindergarten 1-Year Effects V-11

5.4 University of Oregon V-12

  5.4.1 Sample Descriptions V-12

  5.4.2 Kindergarten 1-Year Effects V-12

5.5 University of Kansas, Behavior Analysis Approach V-14

  5.5.1 Sample Descriptions V-14

  5.5.2 Kindergarten 1-Year Effects V-14

5.6 The High/Scope Foundation, Cognitively Oriented Curriculum V-16

  5.6.1 Sample Descriptions V-16

  5.6.2 Kindergarten 1-Year Effects V-18

5.7 University of Florida Parent Education Model V-18

  5.7.1 Sample Descriptions V-18

  5.7.1 Kindergarten 1-Year Effects V-20

5.8 The Educational Development Center (EDC) V-22

  5.8.1 Sample Descriptions V-22

  5.8.2 Kindergarten 1-Year Effects V-22

5.9 University of Pittsburgh, Primary Education Project V-24

  5.9.1 Sample Descriptions V-24

  5.9.2 Kindergarten 1-Year Effects V-25

5.10 Southwest Education Development Laboratory (SEDL) V-25

  5.10.1 Sample Descriptions V-25

  5.10.2 Kindergarten 1-Year Effects V-27

5.11 Conclusions V-29

CHAPTER VI, TWO-YEAR EFFECTS VI-1

6.0 Introduction VI-1

6.1 Far West Laboratory, Responsive Educational Program VI-5

  6.1.1 Overall 2-Year Effects VI-5

  6.1.2 Site Descriptions VI-8

  6.1.3 By Site 2-Year Effects VI-13

6.2 University of Arizona, Tucson Early Educational Model VI-21
  6.2.1 Overall 2-Year Effects VI-21
  6.2.2 Site Descriptions VI-21
  6.2.3 By Site 2-Year Effects VI-23
6.3 Bank Street College Approach VI-35
  6.2.1 Overall 2-Year Effects VI-35
  6.2.2 Site Descriptions VI-38
  6.2.3 By Site 2-Year Effects VI-38
6.4 University of Oregon VI-44
  6.4.1 Overall 2-Year Effects VI-44
  6.4.2 Site Descriptions VI-44
  6.4.3 By Site 2-Year Effects VI-51
6.5 University of Kansas Behavior Analysis Approach VI-56
  6.5.1 Overall 2-Year Effects VI-56
  6.5.2 Site Descriptions VI-56
  6.5.3 By Site 2-Year Effects VI-63
6.6 High/Scope Foundation, Cognitively Oriented Curriculum VI-69
  6.6.1 Overall 2-Year Effects VI-69
  6.6.2 Site Descriptions VI-71
  6.6.3 By Site 2-Year Effects VI-78
6.7 University of Florida Parent Educational Model VI-85
  6.7.1 Overall 2-Year Effects VI-85
  6.7.2 Site Descriptions VI-88
  6.7.3 By Site 2-Year Effects VI-93
6.8 The Educational Development Center (EDC) VI-98
  6.8.1 Overall 2-Year Effects VI-98
  6.8.2 Site Descriptions VI-98
  6.8.3 By Site 2-Year Effects VI-103
6.9 University of Pittsburgh, Primary Education Project VI-108
  6.9.1 Overall 2-Year Effects VI-108
  6.9.2 Site Descriptions VI-112
  6.9.3 By Site 2-Year Effects VI-112
6.10 Southwest Education Development Laboratory (SEDL) VI-118
  6.10.1 Overall 2-Year Effects VI-118
  6.10.2 Site Descriptions VI-122
  6.10.3 By Site 2-Year Effects VI-126
CHAPTER VII, CONCLUSIONS AND RECOMMENDATIONS VII-1

VOLUME IIA

List of Tables

TABLE		PAGE
II-1	Successive Restriction Placed on Ancova Analytic Sample	II-3
II-2	Distribution of Children Meeting Certain Sample-Defining Criteria by Sponsor by Site	II-4
II-3	Number of Pupils Present at Different Exclusion Levels in Ancova Analytic Sample Development	II-6
II-4	Distribution of Final Ancova Sample for Child and School Level Studies by Sponsor by FT/NFT	II-7
II-5	Distribution of By-Site Analytic Sample by FT/NFT	II-9
II-6	Number of Pupils Upon Which Each Reliability Coefficient is Based and a Documentation of Differences Between the Numbers and a Random Sample of 6442	II-10
II-7	Comparative Data on Fall WRAT for a Random Sample and a Traced Longitudinal Random Sample by Sponsor by FT/NFT	II-12
II-8	Comparative Data on PPVT for a Random Sample and a Traced Longitudinal Random Sample by Sponsor by FT/NFT	II-13
II-9	Spring Kindergarten PPVT Means, Standard Deviations and Sample Sizes by Sponsor by FT/NFT for Five Groups of Children	II-9
II-10	Spring Kindergarten MAT Reading Subtest Means, Standard Deviations, and Sample Sizes by Sponsor by FT/NFT for Five Groups of Children	II-10
II-11	Spring Kindergarten MAT Subtest Means, Standard Deviations,	II-11
II-12	Spring First Grade MAT Total Reading Subtest Means, Standard Deviations, and Sample Sizes by Sponsor by FT/NFT for Five Groups of Children	II-12
II-13	Spring First Grade MAT Mathematics Subtest Means, Standard Deviations, and Sample Sizes by Sponsor by FT/NFT for Five Groups of Children	II-13

List of Tables (continued)

TABLE PAGE

II-14 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Far West by FT/NFT II-24

II-15 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Arizona by FT/NFT II-25

II-16 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Bank Street by FT/NFT II-27

II-17 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Oregon by FT/NFT II-28

II-18 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Kansas by FT/NFT II-30

II-19 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for High Scope by FT/NFT II-31

II-20 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Florida by FT/NFT II-32

II-21 Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for EDC by FT/NFT II-34

List of Tables (continued)

TABLE		PAGE
II-22	Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for Pittsburgh by FT/NFT	II-35
II-23	Means and Standard Deviations on Pretest and SES Measures for the population of children tested in Fall 1971 and the 1973 and 1974 analytic sample for SEDL by FT/NFT	II-38
VI-1	Pupil and Family Background Characteristics for Far West, Site NC1	VI-9
VI-2	Pupil and Family Background Characteristics for Far West, Site NC2	VI-10
VI-3	Pupil and Family Background Characteristics for Far West, Site NE	VI-11
VI-4	Pupil and Family Background Characteristics for Far West, Site W1	VI-12
VI-5	Pupil and Family Background Characteristics for Far West, Site W2	VI-14
VI-6	Pupil and Family Background Characteristics for Arizona, Site S	VI-25
VI-7	Pupil and Family Background Characteristics for Arizona, Site NC1	VI-26
VI-8	Pupil and Family Background Characteristics for Arizona, Site NC2	VI-27
VI-9	Pupil and Family Background Characteristics for Arizona, Site NE1	VI-28
VI-10	Pupil and Family Background Characteristics for Arizona, Site NE2	VI-29
VI-11	Pupil and Family Background Characteristics for Bank Street, Site NE1	VI-40

List of Tables (continued)

TABLE		PAGE
VI-12	Pupil and Family Background Characteristics for Bank Street, Site NE2	VI-41
VI-13	Pupil and Family Background Characteristics for Oregon, Site NC1	VI-47
VI-14	Pupil and Family Background Characteristics for Oregon, Site NC2	VI-48
VI-15	Pupil and Family Background Characteristics for Oregon, Site NE	VI-50
VI-16	Pupil and Family Background Characteristics for Kansas, Site NE	VI-58
VI-17	Pupil and Family Background Characteristics for Kansas, Site NC2	VI-59
VI-18	Pupil and Family Background Characteristics for Kansas, Site S	VI-60
VI-19	Pupil and Family Background Characteristics for Kansas, NC1	VI-61
VI-20	Pupil and Family Background Characteristics for High/Scope, Site NE	VI-72
VI-21	Pupil and Family Background Characteristics for High/Scope, Site NC	VI-73
VI-22	Pupil and Family Background Characteristics for High/Scope, Site W1	VI-74
VI-23	Pupil and Family Background Characteristics for High/Scope, Site W2	VI-75
VI-24	Pupil and Family Background Characteristics for High/Scope, Site S	VI-76
VI-25	Pupil and Family Background Characteristics for High/Scope, Site W3	VI-77
VI-26	Pupil and Family Background Characteristics for Florida, Site S1	VI-89
VI-27	Pupil and Family Background Characteristics for Florida, Site S2	VI-90

List of Tables (continued)

TABLE		PAGE
VI-28	Pupil and Family Background Characteristics for Florida, Site W	VI-91
VI-29	Pupil and Family Background Characteristics for Florida, Site NC	VI-92
VI-30	Pupil and Family Background Characteristics for EDC, Site NE1	VI-100
VI-31	Pupil and Family Background Characteristics for EDC, Site NE2	VI-101
VI-32	Pupil and Family Background Characteristics for EDC, Site S	VI-102
VI-33	Pupil and Family Background Characteristics for Pittsburgh, Site NC	VI-113
VI-34	Pupil and Family Background Characteristics for Pittsburgh, Site NE	VI-114
VI-35	Pupil and Family Background Characteristics for SEDL, Site NE	VI-123
VI-36	Pupil and Family Background Characteristics for SEDL, Site W	VI-124
VI-37	Pupil and Family Background Characteristics for SEDL, Site S	VI-125

VOLUME IIA

List of Figures

FIGURE		PAGE
V-1	Kindergarten Effects Profile for Far West on Child Level ANCOVA Analytic Sample	V-7
V-2	Kindergarten Effects Profile for Arizona on Child Level ANCOVA Analytic Sample	V-10
V-3	Kindergarten Effects Profile for Bank Street on Child Level ANCOVA Analytic Sample	V-13
V-4	Kindergarten Effects Profile for Oregon on Child Level ANCOVA Analytic Sample	V-15
V-5	Kindergarten Effects Profile for Kansas on Child Level ANCOVA Analytic Sample	V-17
V-6	Kindergarten Effects Profile for High Scope on Child Level ANCOVA Analytic Sample	V-19
V-7	Kindergarten Effects Profile for Florida on Child Level ANCOVA Analytic Sample	V-21
V-8	Kindergarten Effects Profile for EDC on Child Level ANCOVA Analytic Sample	V-23
V-9	Kindergarten Effects Profile for Pittsburgh on Child Level ANCOVA Analytic Sample	V-26
V-10	Kindergarten Effects Profile for SEDL on Child Level ANCOVA Analytic Sample	V-28
VI-1	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Far West	VI-6
VI-2	Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: Far West - Outcome: MAT Total Reading Subtest	VI-7
VI-3	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Far West, Site NC1	VI-15
VI-4	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Far West, Site NC2	VI-16

List of Figures (continued)

FIGURE PAGE

VI-5 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Far West, Site NE VI-18

VI-6 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Far West, Site W1 VI-19

VI-7 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Far West, Site W2 VI-20

VI-8 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Arizona VI-22

VI-9 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Arizona, Site S VI-30

VI-10 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Arizona, Site NC1 VI-31

VI-11 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Arizona, Site NC2 VI-32

VI-12 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Arizona, Site NE1 VI-33

VI-13 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Arizona, Site NE2 VI-34

VI-14 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample Bank Street VI-36

VI-15 Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: Bank Street - Outcome: Total Reading Subtest VI-37

VI-16 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Bank Street, Site NE1 VI-42

VI-17 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Bank Street, Site NE2 VI-43

VI-18 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Oregon VI-45

VI-19 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Oregon, Site NC1 VI-53

List of Figures (continued)

FIGURE		PAGE
VI-20	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Oregon, Site NC2	VI-54
VI-21	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Oregon, Site NE	VI-55
VI-22	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Kansas	VI-57
VI-23	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Kansas, Site NE	VI-64
VI-24	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Kansas, Site S	VI-66
VI-26	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Kansas, Site NC1	VI-67
VI-27	End of First Grade FT/NFT Contrasts Traced 2-Year Sample, High/Scope	VI-70
VI-28	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, High/Scope, Site NE	VI-79
VI-29	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, High/Scope, Site NC	VI-80
VI-30	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, High/Scope, Site W1	VI-81
VI-31	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, High/Scope, Site W2	VI-82
VI-32	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, High/Scope, Site S	VI-83
VI-33	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, High/Scope, Site W3	VI-84
VI-34	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Florida	VI-86
VI-35	Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: Florida - Outcome: MAT Mathematics Subtest	VI-87
VI-36	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Florida, Site S1	VI-94

List of Figures (continued)

FIGURE		PAGE
VI-37	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Florida, Site S2	VI-95
VI-38	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Florida, Site W	VI-96
VI-39	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Florida, Site NC	VI-97
VI-40	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, EDC	VI-99
VI-41	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, EDC, Site NE1	VI-104
VI-42	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, EDC, Site NE2	VI-105
VI-43	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, EDC, Site S	VI-106
VI-44	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Pittsburgh	VI-109
VI-45	Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: Pittsburgh - Outcome: MAT Total Reading Subtest	VI-110
VI-46	Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: Pittsburgh - Outcome: MAT Mathematics Subtest	VI-111
VI-47	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Pittsburgh, Site NC	VI-115
VI-48	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, Pittsburgh, Site NE	VI-116
VI-49	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, SEDL	VI-119
VI-50	Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: SEDL - Outcome: Mathematics Subtest	VI-120
VI-51	Heterogeneous Regressions: Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT Sponsor: SEDL - Outcome: MAT Total Reading Subtest	VI-121
VI-52	End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, SEDL, Site NE	VI-127

List of Figures (continued)

FIGURE PAGE

VI-53 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, SEDL, Site W VI-128

VI-54 End of First Grade FT/NFT Contrasts, Traced 2-Year Sample, SEDL, Site S VI-129

APPENDICES

CONTENTS

APPENDIX A: THE ANALYTIC SAMPLE

1.0 Introduction A-1

1.1 Pupil Characteristics A-3

1.2 Family Background Characteristics A-11

1.3 Teacher Characteristics A-19

1.4 Classroom Characteristics A-25

1.5 Entering Pretest Characteristics of Classes A-31

1.6 School Characteristics A-32

1.7 Community Characteristics A-42

APPENDIX B: METHODOLOGICAL ISSUES

1.0 Introduction B-1

2.0 The General Linear Relationship Model B-2

- 2.1 Linearity Assumptions B-2
- 2.2 Correlations: Partial, Multiple, and Semi-partial B-3
- 2.3 Three Variations of the Linear Model B-7

3.0 The Analysis of Covariance B-10

- 3.1 The Noise and the Signals B-10
- 3.2 The Covariate Set B-13
- 3.3 The Desirability of a Covariate Set B-15
- 3.4 Comparison of Means B-17
- 3.5 Study of Interactions B-18

4.0 Analytic Method B-21

- 4.1 Missing Data B-21
- 4.2 Coding of Nominal Variables B-21
- 4.3 Exploration of Homogeneity (Heterogeneity of Regression) B-22
- 4.4 Determination of Unadjusted and Adjusted Effects B-22

APPENDICES

CONTENTS (continued)

APPENDIX C: SITE LEVEL KINDERGARTEN EFFECTS PROFILE TABLES C-1

APPENDIX D: DESCRIPTION OF THE FOLLOW THROUGH SPONSORS

1.0 Introduction D-1
2.0 Far West Laboratory D-2
3.0 University of Arizona D-3
4.0 Bank Street College D-4
5.0 University of Oregon D-5
6.0 University of Kansas D-6
7.0 High/Scope Educational Research Foundation D-7
8.0 University of Florida D-8
9.0 Educational Development Center D-9
10.0 University of Pittsburgh D-10
11.0 Southwest Educational Development Laboratory D-11

APPENDICES

List of Tables

Table Page

A-1 Means and Standard Deviations of Child's Age in Months by Sponsor by FT/NFT for ANCOVA Analytic Sample A-4

A-2 Sex of Child by Sponsor by FT/NFT for ANCOVA Analytic Sample A-5

A-3 Student Ethnicity by Sponsor by FT/NFT for ANCOVA Analytic Sample A-6

A-4 Previous Preschool Experience by Sponsor by FT/NFT A-8

A-5 Means and Standard Deviations of Fall Kindergarten WRAT by Sponsor by FT/NFT A-10

A-6 Means and Standard Deviations of Fall PPVT by Sponsor by FT/NFT for ANCOVA Analytic Sample A-12

A-7 Means and Standard Deviations of Family Income Index by Sponsor by FT/NFT for ANCOVA Analytic Sample A-13

A-8 Means and Standard Deviations of Mother's Education Index by Sponsor by FT/NFT for ANCOVA Analytic Sample A-14

A-9 Family Structure by Sponsor by FT/NFT for ANCOVA Analytic Sample A-15

A-10 Means and Standard Deviations of Number of Persons in Household by Sponsor by FT/NFT for ANCOVA Analytic Sample A-16

A-11 Means and Standard Deviations of Parent General Locus of Control by Sponsor by FT/NFT for ANCOVA Analytic Sample A-18

A-12 Means and Standard Deviations of Parent Perception of School Receptivity by Sponsor by FT/NFT for ANCOVA Analytic Sample A-20

A-13 Means and Standard Deviations of Level of Kindergarten Teacher's Education by Sponsor by FT/NFT for ANCOVA Analytic Sample A-21

A-14 Means and Standard Deviations of Level of First Grade Teacher's Education by Sponsor by FT/NFT for ANCOVA Analytic Sample A-23

APPENDICES

List of Tables (continued)

Table		Page
A-15	Means and Standard Deviations of First Grade Teacher's Elementary Teaching Experience in Years by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-24
A-16	Kindergarten Teacher Ethnicity by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-26
A-17	First Grade Teacher Ethnicity by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-27
A-18	Kindergarten Classroom Ethnicity by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-28
A-19	First Grade Classroom Ethnicity by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-29
A-20	Means and Standard Deviation of Class Mean Fall Kindergarten WRAT by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-31
A-21	School Ethnicity by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-33
A-22	Title I Funds per Pupil in School by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-34
A-23	Means and Standard Deviations of Man Days of Professional Services per Pupil per Week by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-35
A-24	Means and Standard Deviations of Man Days of Instructional Services per Pupil per Week by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-36
A-25	Man Days of Social Services per Pupil per Week by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-37
A-26	Man Days of Medical Services per Pupil per Week by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-38
A-27	Means and Standard Deviations of Total Enrollment in School by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-40
A-28	Means and Standard Deviations of Student-Teacher Ratio in School by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-41
A-29	Regional Distribution of Children by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-43

APPENDICES

List of Tables (continued)

Table		Page
A-30	Distribution of Children by City Size by Sponsor by FT/NFT for ANCOVA Analytic Sample	A-44
B-1	F Ratios and Factors in the Hierarchical Model	B-23

APPENDICES

List of Figures

Figure		Page
C -1	Kindergarten Effects Profile for Far West Site NC2	C -1
C -2	Kindergarten Effects Profile for Far West Site NC1	C -2
C -3	Kindergarten Effects Profile for Far West Site W1	C -3
C -4	Kindergarten Effects Profile for Far West NE	C -4
C -5	Kindergarten Effects Profile for Far West Site W2	C -5
C -6	Kindergarten Effects Profile for Arizona Site S	C -6
C -7	Kindergarten Effects Profile for Arizona Site NC1	C -7
C -8	Kindergarten Effects Profile for Arizona Site NC 2	C -8
C -9	Kindergarten Effects Profile for Arizona Site NE1	C -9
C -10	Kindergarten Effects Profile for Arizona Site NE2	C -10
C -11	Kindergarten Effects Profile for Bank Street Site NE2	C -11
C -12	Kindergarten Effects Profile for Bank Street Site NE1	C -12
C -13	Kindergarten Effects Profile for Oregon Site NC1	C -13
C -14	Kindergarten Effects Profile for Oregon Site NC2	C -14
C -15	Kindergarten Effects Profile for Oregon Site NE	C -15
C -16	Kindergarten Effects Profile for Kansas Site NE	C -16
C -17	Kindergarten Profile for Kansas Site NC2	C -17
C -18	Kindergarten Effects Profile for Kansas Site S	C -18
C -19	Kindergarten Effects Profile for Sansas Site NC1	C -19
C -20	Kindergarten Effects Profile for High/Scope Site NE	C -20
C -21	Kindergarten Effects Profile for High/Scope Site NC	C -21
C -22	Kindergarten Effects Profile for High/Scope Site W1	C -22

APPENDICES

List of Figures (continued)

Figure		Page
C -23	Kindergarten Effects Profile for High/Scope Site W2	C -23
C -24	Kindergarten Effects Profile for High/Scope Site S	C -24
C -25	Kindergarten Effects Profile for High/Scope Site W3	C -25
C -26	Kindergarten Effects Profile for Florida Site S1	C -26
C -27	Kindergarten Effects Profile for Florida Site S2	C -27
C -28	Kindergarten Effects Profile for Florida Site W	C -28
C -29	Kindergarten Effects Profile for Florida Site NC	C -29
C -30	Kindergarten Effects Profile for EDC Site NE1	C -30
C -31	Kindergarten Effects Profile for EDC Site NE2	C -31
C -32	Kindergarten Effects Profile for EDC Site S	C -32
C -33	Kindergarten Effects Profile for Pittsburgh Site NC	C -33
C -34	Kindergarten Effects Profile for Pittsburgh Site NE	C -34
C -35	Kindergarten Effects Profile for SEDL Site NE	C -35
C -36	Kindergarten Effects Profile for SEDL Site W	C -36
C -37	Kindergarten Effects Profile for SEDL Site S	C -37

Preface

This is the second annual report on the National Evaluation of Follow Through being carried out by Abt Associates Inc., under Contract #OEC-0-72-5221 with the United States Office of Education. The first report, "Education as Experimentation: Early Effects of Follow Through" described the program and the history of the evaluation effort. It then reported the findings of the comparisons between the performances of the Follow Through and Non-Follow Through children of Cohort III (the largest and most stable of the Planned Variation Cohorts) at the end of their kindergarten year. This second report focuses on a subset of those kindergarteners: those who completed first grade with the same Sponsor and for whom end-of-first grade test scores were available. Thus, this report is a direct continuation of the first volume and should be read with the contents of Volume I in mind. This report may, however, be read independently as we have tried to provide within it that information from the first annual report which is prerequisite to an understanding of this year's data.

The primary goal of this volume, "Two Year Effects of Follow Through," is to describe the comparisons between Follow Through (FT) and Non-Follow Through (NFT) children at the end of first grade. To accomplish this analysis, only children who have experienced this two year treatment have been entered into the studies. As a result of this decision, we are now dealing with a different subset of children than those analyzed in the first kindergarten study. This report, therefore, considers the differences, on demographic and entry level characteristics, between the full kindergarten sample and the two year longitudinal sample. The reader should keep these differences firmly in mind when considering the findings at the end of first grade. In order to facilitate these considerations, the analysis applied to the original kindergarten group has been applied to the longitudinal subset and reported here.

Volume IIB, "Monograph s and Appendices" contains a great deal of information with which the technical reader is concerned. Much of the detail or samples, instruments, and analytic procedures has been placed in the Appendices in order to increase the usefulness of Volume IIA. With the findings reported in IIA, we have included the least amount of such technical information necessary to provide a technical background for the reader. We hope that the decision rules used to divide technical information between Volumes IIA and IIB serve the purpose for which they are intended: to make the findings clear and to allow the reader the latitude for critical reading which we believe the findings deserve.

Acknowledgements

Essentially the same staff and consultants participated in the work that produced this Volume as those who helped with the first report. Dr. Jacob Cohen of New York University served as our chief methodological consultant. Dr. William Meyer of Syracuse University served as our External Reviewer and contributed a wise and useful perspective to our work.

The remaining panel of consultants included Dr. E. Kuno Beller, Temple University; Dr. David C. Berliner, Far West Regional Laboratory; Dr. Robert Brennan, SUNY at Stony Brook; Dr. Jimmie C. Fortune, Virginia Polytechnic Institute; Dr. Edward J. O'Connell, Syracuse University. The panel have met with us both as a group and individually and they have provided us with criticisms and encouragement that we have come to expect from such an active and highly professional group of colleagues.

We would also like to acknowledge the very helpful suggestions of Mr. Eugene Tucker, Project Officer and Director of the Office of Evaluation of Follow Through in O.P.B.E. and Dr. Mary Kennedy of the Follow Through evaluation staff.

CHAPTER I

OVERVIEW

1.0 INTRODUCTION

This is the second annual report on the National Evaluation of Follow Through being carried out by Abt Associates Inc., under Contract #OEC-0-72-5221 with the United States Office of Education. The first report, "Education as Experimentation: Early Effects of Follow Through", included a series of studies of kindergarten year of Cohort III children who attended Follow Through classes under the aegis of ten Follow Through Sponsors.[1] A comprehensive description of Follow Through, the rationale of the Planned Variation approach, and a discussion of many of the issues involved in the analysis of a large, complex data base are to be found in the first Report, and will not be repeated here. We remind the reader, however, that the primary focus of these studies is on Cohort III, because here we find the largest and most stable sample of children. Furthermore, each Sponsor included in these studies has had at least two years' experience in administering its model prior to operating with Cohort III. The children included in the earlier Cohorts will be used to examine some issues associated with the increasing maturity of the Sponsors in delivering their models, but these studies are not included in the present report.

This second report differs from the first in two major respects. First, this report focuses on a group of Cohort III children who have been traced from kindergarten to the first grade. These are children who remained in the same schools with the same Sponsor and were tested in both years. These children may differ in important ways from children who do not remain in the program, or who did not get included in the tested sample at kindergarten or first grade. There is evidence that some systematic differences exist between those who continue and those who do not continue in the program so it will be necessary for us to examine the nature of the traced sample carefully in order to account for this attrition bias. The advantage of focusing on this group, however, is that they provide an opportunity to identify one and two year effects of Follow Through with an identical analytic model and sample.

---

[1] The ten Sponsors of interest are described in Appendix D of this volume.

The second important difference between this report and the first is that we shall add site level analyses to the previous child and school level analyses. This is particularly important since the first report indicated that many of the Sponsor effects were strongly influenced by the specific conditions which are to be found at each site. Two questions then arose. The first is the extent to which Sponsors can be compared to each other if they are located at different sites (and except for a very small group of Sponsors who are operating in the same large cities, there is no overlap of Sponsor location). The second is the degree to which each Sponsor's effects are differentially influenced by the several sites in which the program is operating. Some Sponsors have very site-specific effects, whereas others seem able to overcome the differences in sites to produce uniform results everywhere. This kind of finding is a central consideration in policy decision making, of course, and we shall consider the issue in this Report along with the major task of identifying Sponsor Follow Through (FT)/Non Follow Through (NFT) contrasts.

In order to describe the contents of this report, we now summarize the general analytic strategy we have adopted for the series of studies. The earlier report presents a fuller description of this strategy.

1.1 GENERAL ANALYTIC APPROACH

Our primary goal is to identify and describe the effects of each Sponsor's program upon children with various characteristics, at various grade levels, and under various conditions of program administration. If this study is to contribute to policy considerations, we must document the range of impacts produced by each of the educational models, so that the Federal Government can provide local educational agencies with the information they need to make choices about means to achieve locally defined goals. An informed choice, in this context, means that the local agency will have a reasonable estimate of the likely effects produced by each of the models under the particular conditions found at the local site. To accomplish this goal unambiguously, each Sponsor would have to administer its model within the same (or a similar) range of local conditions. We would then have data from which we could create a directory of relative effects which would provide the local consumer with the resources for an informed decision.

Sponsors were not assigned to comparable sets of sites. Each Sponsor's distinctive way of working in school systems, moreover, necessarily reduces the comparability of the different sites within which the different Sponsors are operating. Differences in the types of children and the types of areas in which Sponsors are applying their models make direct comparisons among Sponsors difficult. The problem is further confounded by differences in the relationships between site and child characteristics.

We first asked this question with respect to the kindergarten year of Cohort III; the answers are reported in the first volume of this series. For most Sponsors, we found some important contrasts between Follow Through and Non Follow Through groups. For some cases, the contrasts were in the academic domain, and for some, in the affective domain. A few Sponsors showed no important contrasts at the end of the kindergarten year, although some trends may have been developing in the achievement area. Note that these contrasts were ascribed, within the limits of a small set of adjusting variables, to the overall presence of the Sponsors in the lives of the children, and not to anything in particular which the Sponsors were doing in the classroom or in the school system. These contrasts emanate from black boxes, and we are in no position to attribute them unambiguously to the curriculum (i.e., the model), to the strategy of change adopted by the Sponsor, (i.e., the program), or to the unique interaction between these two at any site for any Sponsor.

In order to be useful for policy decisions, this phenomenon of a variety of important FT/NFT contrasts scattered across the several Sponsors needs to be robust. It would make no sense to build policy upon a set of findings which were not stable at least over time and, in a special sense, over grade levels. Stability over time means that the kindergarten effect needs to be replicated in several kindergarten years. In the Follow Through evaluation, this is difficult to accomplish despite the presence of several cohorts. The first two of these cohorts represented developmental periods for the Sponsors, and involved somewhat restricted samples of children. Stability over time is thus confounded with changing historical events. Nevertheless, trends can be discerned in such comparisons of effects across cohorts and these trends can provide heuristic assistance to the policy decisions.

Stability over grade levels is equally important, but somewhat more complex than stability over time. Despite the fact that the children participating in a longitudinal study over grade are the same people throughout the study, they are very different kinds of people at each grade level from the point of view of an educational study. First grade children are considerably more mature than kindergartners, and can be expected to have rather different needs in both the intellectual and affective domains. There is no reason to expect that all effects in the kindergarten year should be found again in the first grade. Indeed, these data appear to document multiple routes to academic growth. Stability over grade levels must mean, therefore, stability of a pattern of effects in the same children involved with the same Sponsor over grade span. We can expect that each Sponsor will have a unique kindergarten effect that is observable each time the Sponsor administers a kindergarten program. We can also expect the Sponsor to have a unique first grade effect, and a unique second and third grade effect each time these grades are taught. This is stability over time, and it is critical to know how that Sponsor interacts with children at each developmental level.

The cumulative effect over grades is still another concern in a search for stability over grades. We would want to know if the effects of one year are important to the production of effects the next year. Even more interesting, we would want to know what kinds of effects in the kindergarten year are important to what kinds of effects in the first grade, and so on throughout the full Follow Through grade span. In this manner we maximize precisely the information which each local educational decision maker needs to have in order to make the most reasonable choice for his site.

Next, from the point of view of national policy making, we must examine the effects which are found to be important and stable for their generalizability across places, people, and conditions. Each Sponsor has a moderate distribution over these parameters, but in order to feel confident that the effects can be generalized beyond the Follow Through groups, it is necessary to examine them as they relate to the variations in site location, kinds of children, and conditions of administration. To the extent that such relationships can be established with confidence, the validity of the effects can be defined for groups external to the Follow Through children and systems. An evaluation design capable of producing these analyses must include the interactions of Sponsors with places, types of children, and conditions of administration.

For a variety of reasons, not all of these issues can be considered in this phase of the national evaluation. Rather, this report will focus on a selected set of issues which include FT/NFT comparisons for each Sponsor using Cohort III children in such a way that:

(a) all tested and traced FT children for a given Sponsor are compared to all tested and traced NFT children for that Sponsor at the end of kindergarten (one year effects).

(b) all tested and traced FT children for a given Sponsor are compared to all tested and traced NFT children for that Sponsor, at the end of first grade (two year effects).

(c) establishment of the stability of the two year effects over the several sites within which each Sponsor is working.

Although it is not a major goal of this report to compare the one and two year effects for a given Sponsor, there is some advantage in having such an analysis available. Thus, a study of cumulative effects over grades is reported as Monograph 1 in Volume IIB of this Report.

Several sets of descriptive materials on teachers and schools are presented in Appendices to this Report. Some of these data have been related to the description of the sample of children also reported in Appendix A of this volume. The reader is referred to the Appendices, and urged to interpret the effects reported in this volume in the light of these descriptive materials.

It should also be noted that the first report in this series included an extensive description of the analytic procedures used with these data. Appendix B of this volume includes selected portions of that description so that the technical reader will not have to refer to the earlier volume to answer any questions which arise about the present analysis. The present

Report includes a description of the analysis plan generated to meet the needs of the two year overall and site level studies in order to emphasize the continuity between the two reports.

CHAPTER II

SAMPLE DEVELOPMENT[1]

2.0 INTRODUCTION

This chapter is divided into three sections.

First we define the various groups of children studied in this report. These groups are the ANCOVA analytic sample, the by-site analytic sample and the reliability samples. A full description of the analytic sample is presented in Appendix A of this volume.

The second section examines several steps of the sample-defining process to determine the extent to which the various selection criteria bias the sample. The question of interest is: how well does the current sample represent the population of FT/NFT students tested in the Fall of 1971? Unfortunately we have no way of assessing differences between the tested population and the general population of all FT/NFT children. Therefore any statement as to the generalizability refers to generalizing about the tested population only.

Within the second section of this chapter we also compare the current sample to the sample analyzed in last year's report. Differences in outcomes across the years may be related to differences in the samples upon which the analyses were performed.

Finally the issue of Sponsor to Sponsor comparisons is briefly addressed in terms of the findings in this chapter and Appendix A of this volume.

---

[1] The word "sample" is used here to refer to any group of children analyzed in this report. The groups of children studied might be more accurately referred to as subsets.

2.1 SAMPLE SELECTION CRITERIA

2.1.1 1974 Analytic Samples

This year's analytic subset is composed of Cohort III children who entered kindergarten in the Fall of 1971. The following series of successive restrictions were placed upon the data in the process of defining our analytic samples.

In order to be a member of Cohort III entering kindergarten, a child must have been present for at least one data collection event during that entering year. Our initial requirement for analytic purposes was that the child had taken the fall kindergarten WRAT. A total of 17,721 pupils met this requirement as indicated in Table II-1. Our second requirement, a complete achievement test battery (i.e. in addition to the Fall 1971 WRAT, the Spring 1972 WRAT and MAT and the Spring 1973 MAT), was met by 4,962 of these pupils. Table II-2 presents the site level breakdown of these pupils. With the additional restriction that children attend single grade classes in the first grade rather than ungraded or mixed grade classes, 4,742 pupils were retained. The requirement that each pupil remain in this original treatment group reduced the N to 4,625, which is also displayed at site level in Table II-2.

At this point, the ANCOVA and by-site sample definition procedures differ. We first discuss the ANCOVA sample definition procedure and later turn to that of the by-site sample.

For the ANCOVA sample, two restrictions were imposed to increase the comparability of pupil and school level analyses. Requiring successively that the pupil remain in the same school and that each school have five or more pupils resulted in reductions of the sample size to 4,467 and 4,418 respectively. School changes involving kindergarten schools feeding into one elementary school or administrative shifts of a large group of children were permitted. As a result, our sample of children will be distributed in a different number of kindergarten and first grade schools.

In order to use a common sample of pupils and a common set of covariates across all dependent measures, the restrictions that a child must have all dependent measures and all covariates were introduced. The dependent measures required, in addition to the aforementioned achievement test battery, were

TABLE II-1

SUCCESSIVE RESTRICTIONS PLACED ON ANCOVA ANALYTIC SAMPLE

Restriction	Resultant N
Fall 1971 WRAT	17721
Spring 1972 WRAT Spring 1972 MAT Spring 1973 MAT	4962
First grade status in 1972-73	4742
Remain in original treatment group	4625
Retention in same school	4467
N ≥ 5 in school	4418
Complete set of dependent measures	3665
Complete set of covariates	2790
N ≥ 5 in school	2755

TABLE II-2

DISTRIBUTION OF CHILDREN MEETING CERTAIN SAMPLE - DEFINING CRITERIA BY SPONSOR BY SITE

Sponsor	Location[a]	NUMBER OF PUPILS With 4 tests[b]	 Who achieved first grade status and remained in same treatment group	 Who are present in this year's analysis
Far West	L-NE	134	130	36
	M-NC	181	163	110
	M-W	87	76	56
	L-NC	107	85	67
	M-W	110	109	94
	M-W	60	56	34
Total		679	619	397
Arizona	S-S	112	95	50
	L-NC	65	40	32
	S-NE	116	114	83
	L-NE	100	98	71
	M-NC	255	243	89
	L-NC	131	126	101
Total		779	716	426
Bank Street	S-NE	64	62	46
	M-NE	133	128	94
	L-NE	46	46	32
	L-NE	57	49	14
Total		300	285	186
Oregon	M-NC	96	87	64
	L-NE	24	24	-
	M-NC	114	113	87
	M-NE	75	63	26
Total		309	287	177
Kansas	L-NE	68	98	29
	L-NE	104	68	78
	M-NC	107	101	93
	M-NE	177	171	58
	L-NC	160	154	93
	L-S	121	116	82
Total		737	708	433
High Scope	S-S	97	97	67
	L-NE	72	72	33
	L-NC	92	89	68
	S-W	59	54	54
	L-W	92	92	57
	L-W	97	95	34
Total		509	499	313
Florida	L-NE	25	23	19
	L-NE	106	104	91
	S-W	112	100	59
	L-S	202	200	43
	R-NC	80	79	35
Total		525	506	247
EDC	S-NE	115	111	85
	L-NE	92	52	45
	M-NE	123	111	45
	L-S	63	56	-
	M-NE	87	87	71
Total		480	417	246
Pittsburgh	S-NE	116	109	70
	R-NC	188	157	143
Total		304	266	213
SEDL	L-NE	99	97	8
	M-W	118	103	35
	R-S	123	122	74
Total		340	322	117
Grand Total		4962	4625	2755

[a] Geographic location: NE: Northeast; NC: North Central; W: West; S: South
Size of City: L: Large; M: Medium; S: Small; R: Rural

[b] Fall 1971 WRAT; Spring 1972 WRAT and MAT; Spring 1973 MAT

the spring kindergarten PPVT, Locus of Control and Gumpgookies tests as well as kindergarten and first grade absence data. Zero test scores were not included in the analyses.[2] The covariates required included the Fall kindergarten PPVT (in addition to the already required Fall WRAT) and school ethnic composition data. Missing data codes were planned for family income and mother's education; however, we did require that all children in the ANCOVA analytic sample have parent interview information. The dependent and covariate related restrictions resulted in reductions of the ANCOVA sample size to 3,665 and 2,790 respectively. The distribution of the sample by Sponsor and FT/NFT are presented in Table II-3.

Finally the number of sample children in each school was checked again to insure that all schools had five or more pupils in the sample. This reinstatement of an earlier restriction created a final ANCOVA sample of 2,755 children. These children came from 201 kindergarten schools and 210 first grade schools. The distributions of children and schools by Sponsor by FT/NFT are presented in Table II-4. Table II-2 displays the distribution of these children by site.

To trace the by-site sample definition process we return to the sample size of 4,625 (see Table II-1). The next set of restrictions applied to the by-site sample was the presence of a complete set of dependent measures. This restriction was explained above in the description of the ANCOVA sample definition process. Three exceptions were made in an effort to retain as large a number of sites as possible: one site lacked sufficient kindergarten absence data; a second, first grade absence data; and a third, Gumpgookies test data. Rather than eliminate all the children in these sites, we opted

---

[2] When a child receives a zero score we feel that the test is not measuring the child's performance ability but rather may be reflecting any one of a number of situations. Perhaps the child cannot communicate in English; maybe he/she is psychologically or physically unable to perform in a testing situation. The one exception to the non-zero test score restriction was for the Locus of Control test. If a child achieved a non-zero score on the Gumpgookies test, we concluded that a zero on the Locus of Control test designated an "external" child.

TABLE II-3

NUMBER OF PUPILS PRESENT AT DIFFERENT EXCLUSION LEVELS IN ANCOVA ANALYTIC SAMPLE DEVELOPMENT

Sponsor	NUMBER OF PUPILS			
	Who have a complete set of non-zero dependent measures		Who have a complete set of non-zero dependent measures and a complete covariate set	
	FT	NFT	FT	NFT
Far West	285	199	223	178
University of Arizona	385	249	260	169
Bank Street	140	86	119	67
University or Oregon	127	91	95	86
University of Kansas	445	132	324	113
High Scope	185	168	168	145
University of Florida	287	112	173	78
EDC	208	119	154	96
University of Pittsburgh	170	63	153	60
SEDL	145	69	77	52
Total	2377	1288	1746	1044

TABLE II-4

DISTRIBUTION OF FINAL ANCOVA SAMPLE FOR CHILD AND SCHOOL LEVEL STUDIES BY SPONSOR BY FT/NFT

Sponsor	Child Level		Kindergarten School Level		1st Grade School Level	
	FT	NFT	FT	NFT	FT	NFT
Far West	219	178	23	15	23	15
Arizona	260	166	18	14	19	15
Bank Street	119	67	10	6	10	6
Oregon	95	82	6	6	7	6
Kansas	320	113	18	9	18	9
High Scope	168	145	8	12	9	14
Florida	173	74	15	5	15	5
EDC	150	96	8	8	8	8
Pittsburgh	153	60	6	5	9	5
SEDL	67	50	4	5	4	5
Total	1724	1031	116	85	122	88

to drop the problem dependent measure and analyze all others for these three sites. No covariate related restriction was necessary since the Fall WRAT requirement had previously been met and missing data codes were to be utilized for the only other covariate, mother's education.

The final requirement was that each site have a minimum of 16 pupils in each treatment group. The by-site analytic sample contains 2,168 FT pupils and 1,313 NFT pupils in 37 sites. The sample distribution is displayed in Table II-5. [3]

2.1.2 Reliability Samples

In addition to the ANCOVA and by-site sample definition procedures, a third was utilized to generate a subset of children for determining reliability coefficients for the entire test battery. Here we did not use a traced longitudinal sample--but rather a cross-sectional one. In other words, in contrast to the above procedures we did not require children to have had a complete longitudinal achievement battery. Rather we selected a random sample (N=6,442) of all children on the Cohort III kindergarten data base who had Fall WRAT test data. Then a reliability coefficient was computed for each test using all pupils in this random sample who obtained a score greater than zero on that test. The specific number of pupils available for each reliability coefficient is presented in Table II-6. In order to determine if the tests were functioning similarly in the population and our analytic sample, reliability coefficients were also computed for the 2,755 pupils in the analytic sample.

2.3 REPRESENTATIVENESS OF 1974 ANALYTIC SAMPLE

A number of restrictive criteria have been utilized in developing the analytic samples. In order to determine the extent and direction of sample biases related to the application of these selections of criteria, we have examined psychometric properties of tests for various subsets of children at different stages in the sample definition process. We have also compared the present Cohort III analytic sample to the Cohort III group tested in the Fall 1971 and the Cohort III group analyzed in the 1973 kindergarten study on pretest and SES measures. The first comparison is to ascertain representativeness, the latter to explore comparability of results across the two years. For each of the above purposes we have used the ANCOVA analytic sample.

---

[3] A full description of site characteristics is presented in Chapter VI of this volume.

TABLE II-5

DISTRIBUTION OF BY-SITE ANALYTIC SAMPLE BY FT/NFT

Sponsor	Site[a]	Sample Size	
		FT	NFT
FAR WEST	NE	65	17
	NC1	73	51
	W1	33	27
	NC2	41	35
	W2	51	57
ARIZONA	S	39	46
	NE1	64	27
	NE2	66	20
	NC1	140	92
	NC2	58	57
BANK STREET	NE1	31	27
	NE2	71	51
OREGON	NC1	47	33
	NC2	49	47
	NE	36	20
KANSAS	NC1	58	40
	NE	109	24
	NC_2	97	38
	S	68	29
HIGH SCOPE	S	52	60
	NE	35	17
	W1	36	18
	W_2	34	40
	NC	40	33
	W3	43	34
FLORIDA	S1	61	38
	W	66	33
	S2	102	35
	NC	41	35
EDC	NE1	69	39
	S	24	16
	NE2	45	29
PITTSBURGH	NE	66	30
	NC	118	34
SEDL	NE	36	18
	W	40	40
	S	64	26
TOTAL		2,168	1,313

[a] Site Location: NE = Northeast
NC = North Central
W = West
S = South

TABLE II-6

NUMBER OF PUPILS UPON WHICH EACH RELIABILITY COEFFICIENT IS BASED AND A DOCUMENTATION OF DIFFERENCES BETWEEN THE NUMBERS AND A RANDOM SAMPLE OF 6442

		Tested			
	Did Not Take Test	Obtained Zero Score	Obtained Greater Than Zero Score	Sample Size for Reliability Coefficient	Number of Pupils Obtaining at Least One Subtest Zero Score on Respective Instruments
Pre-School Inventory	106	51	6285	6285	-
Lee-Clark	17	102	6323	6323	-
PPVT (Fall)	146	21	6275	2682	-
WRAT - Spelling (Fall)	-	954	5488	6401	1040
- Reading (Fall)	-	120	6322		
- Arithmetic (Fall)	-	125	6317		
- Total (Fall)	-	41	6401		
MAT - Listening to Sounds (Primer)	3185	89	3168	2996	261
- Reading (Primer)		112	3145		
- Numbers (Primer)		131	3126		
PPVT - (Spring)	3706	29	2707	2682	-
WRAT - Spelling (Spring)	3724	32	2686	2680	38
- Reading (Spring)		3	2715		
- Arithmetic (Spring)		6	2712		
- Total (Spring)		1	2717		
Locus of Control - Positive	3742	94	2606	2695	5
- Negative	3742	105	2595		
Gumpgookies - Instrumental Activities	3754	8	2680	2677	11
- School Enjoyment		6	2682		
- Self-Evaluation		8	2680		
- Purposive Behavior		5	2683		
- Self-Confidence		7	2681		
- Total		3	2685		
MAT - Word Knowledge (Primary I)	3381	34	3027	2938	133
- Word Analysis (Primary I)		51	3010		
- Reading (Primary I)		25	3036		
- Total Reading (Primary I)		3	3058		
- Math Concepts (Primary I)		85	2976		
- Math Computation (Primary I)		179	2882		
- Math Total (Primary I)		50	3011		

NOTE: For comparability purposes, the coefficients for all subtests within a given test were based on a common sample. The inclusion criteria are basically similar to those utilized in the analytic sample. A score greater than zero on the Preschool Inventory, the Lee-Clark, and the total Fall WRAT was required. On the Spring WRAT, the two MATs, and the Gumpgookies, a score greater than zero on each of the subtests was required. On each of the two Locus of Control subtests, a score greater than zero on each of the subtests or a score greater than zero on the Gumpgookies was required. For the PPVT test-retest coefficient, a score greater than zero at each testing time was required.

The following data are presented for those interested in a more detailed breakdown of differences in N's between the random sample of 644s and the N upon which each coefficient is based.

Testing Time	Test	No. of Subtests on which a Zero Score was Obtained	No. of Pupils Obtaining a Zero Score
Fall 1971	WRAT	3	41
		2	77
		1	922
Spring 1972	MAT	3	4
		2	63
		1	194
	WRAT	4	1
		2	1
		1	36
	Gumpgookies	5	3
		4	2
		3	2
		2	1
		1	3
	Locus of Control	2	3
		1	2
Spring 1973	MAT	5	1
		4	2
		3	7
		2	26
		1	97

2.3.1 Comparison of Traced Longitudinal Random Sample with Random Sample of Fall 1971 Tested Population

This first look at the sample defining process explores the longitudinal restriction of data availability at all three testing points (Fall '71, Spring '72 and '73). To do so we turn to the random sample of the Fall 1971 tested population which was employed in determining reliability estimates. Using Fall WRAT (Total) and PPVT scores we compare the subset of pupils in this random sample who have data at all three test points (the traced longitudinal random sample) with the total random sample which was tested in the Fall of 1971. Considering the total random sample as the best estimate of the "population" of interest, we used two heuristic devices for assessing the representativeness of the traced longitudinal sample. The criterion of a sample mean falling within two standard errors of the "population" mean was utilized first. Then, we also noted whether or not the difference between means was as large as a quarter of a standard deviation. Our findings are presented in Tables II-7 and II-8. In the following discussion major emphasis will be placed on those Sponsor FT/NFT groups which satisfy both heuristic criteria.

Whereas the longitudinal NFT sample for the University of Florida has a higher PPVT mean than its "population", the retained FT sample has a slightly higher mean on the WRAT. The increase in the NFT mean on the PPVT makes the FT and NFT longitudinal samples more comparable. However, the increase in the FT mean on the WRAT enlarged the disparity between FT and NFT samples on that instrument.

For the University of Pittsburgh, both the FT and NFT longitudinal samples have higher means than their respective "population" on the PPVT and the WRAT. For this particular Sponsor, those children (both FT and NFT) who were tested at three different points in time have higher scores than those who were not.

SEDL's longitudinal NFT sample has a higher mean on the PPVT than its "population"; SEDL's longitudinal FT sample has a lower mean on the WRAT than its "population". Hence the FT/NFT groups which appear to have started out fairly comparable in the Fall of 1971 have become less comparable when we require longitudinal data with the longitudinal NFT group having a higher entry level.

TABLE II-7

COMPARATIVE DATA ON FALL WRAT[a] FOR A RANDOM SAMPLE AND A TRACED LONGITUDINAL RANDOM SAMPLE BY SPONSOR BY FT/NFT

		Random Sample[b]		Traced Longitudinal Random Sample[c]		
		$\bar{X}$	N	$\bar{X}$	S.D.	N
FAR WEST LABS	FT	26.95	287	29.50	12.18	156
	NFT	27.85	283	29.66	12.05	94
UNIVERSITY OF ARIZONA	FT	28.58	310	28.10	12.34	171
	NFT	32.55	262	31.68	14.69	100
BANK STREET	FT	27.32	240	26.74	12.88	82
	NFT	27.80	257	25.38	12.77	60
UNIVERSITY OF OREGON	FT	28.93	238	29.62	10.20	99
	NFT	29.79	228	29.84	12.03	49
UNIVERSITY OF KANSAS	FT	27.89	398	27.13	10.83	214
	NFT	24.98	314	25.47	13.10	87
HIGH SCOPE FOUNDATION	FT	25.45	175	26.51	12.11	113
	NFT	30.79	164	31.30	14.38	87
UNIVERSITY OF FLORIDA	FT[1]	26.74	231	28.61	12.67	124
	NFT	24.83	193	23.45	10.51	65
EDUCATIONAL DEVELOPMENT	FT	28.21	229	27.02	13.24	93
	NFT	29.65	200	29.52	12.01	66
UNIVERSITY OF PITTSBURGH	FT[1]	31.43	129	34.29	13.17	69
	NFT[2]	30.47	172	36.41	10.84	17
SEDL	FT[2]	25.78	332	20.46	10.82	94
	NFT	24.13	206	22.86	9.35	43

[a] Child must have taken WRAT, but may have had a zero score.

[b] The random sample is drawn from the population of children tested in the Fall of 1971.

[c] The traced longitudinal random sample includes all those children in the random sample who were also present for testing in Spring, 1972 and 1973.

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

TABLE II-8

COMPARATIVE DATA ON PPVT[a] FOR A RANDOM SAMPLE AND A TRACED LONGITUDINAL RANDOM SAMPLE BY SPONSOR BY FT/NFT

		Random Sample[b]		Traced Longitudinal[c] Random Sample		
		$\bar{X}$	N	$\bar{X}$	S.D.	N
FAR WEST LABS	FT	46.07	284	46.33	10.97	156
	NFT	45.31	277	47.08	11.38	93
UNIVERSITY OF ARIZONA	FT	44.63	308	43.96	12.00	170
	NFT	48.56	260	49.71	12.64	99
BANK STREET	FT	40.69	233	40.21	12.51	80
	NFT	41.89	247	40.32	14.28	56
UNIVERSITY OF OREGON	FT	43.81	236	44.21	9.42	99
	NFT	43.49	221	42.71	11.51	48
UNIVERSITY OF KANSAS	FT	39.25	392	39.15	10.83	212
	NFT[1]	37.39	306	40.21	12.34	84
HIGH SCOPE FOUNDATION	FT	39.97	173	40.25	11.91	113
	NFT	43.56	160	44.16	12.30	87
UNIVERSITY OF FLORIDA	FT	44.04	225	45.22	10.22	120
	NFT[2]	40.58	185	43.75	11.90	61
EDUCATIONAL DEVELOPMENT	FT	43.54	226	42.64	10.22	92
	NFT	44.20	196	45.26	11.90	65
UNIVERSITY OF PITTSBURGH	FT[2]	49.04	127	52.07	9.28	69
	NFT[2]	48.13	171	52.00	6.28	17
SEDL	FT	35.72	314	35.15	14.44	87
	NFT[2]	36.33	201	42.34	13.97	41

[a] Child must have taken PPVT, but may have had a zero score.

[b] The random sample is drawn from the population of children tested in the Fall of 1971.

[c] The traced longitudinal random sample includes all those children in the random sample who were also present for testing in Spring, 1972 and 1973.

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

The differential patterns of sample changes noted above may have a number of sources. For example, the testing program may have been systematically discontinued at different types of locations for different groups; different types of children in different groups may have moved thereby leaving the programs; or different types of children may have been systematically removed from the program for administrative or educational reasons. For example, high achieving children in the SEDL model may have been transferred out of the program after kindergarten, whereas lower achieving children may have been transferred out of the Pittsburgh classes. Whatever the reasons, however, these findings suggest that for various Sponsor/treatment groups differential attrition does exist as we generate a traced longitudinal sample. Our first step in the sample defining process seriously complicates our ability to generalize our findings.

2.3.2 Comparison of 1974 Analytic Sample with Traced Longitudinal Sample

In utilizing our random sample of the Fall 1971 tested population we were able to look at the effect on representativeness of requiring data at all three testing points. Now we wish to explore the influence of additional restrictions which we utilized in the development of the analytic sample. We shall explore this issue through an examination of five dependent measures (kindergarten Spring PPVT, kindergarten MAT Reading and Numbers subtests and first grade MAT Total Reading and Mathematics subtests) for five overlapping groups of pupils. The first group consists of the 4,418 pupils who were retained in the ANCOVA sample just prior to the restriction that a pupil must have non-zero scores (the traced longitudinal sample). The second set of data draws from the same group of 4,418 pupils, but excludes those with zero scores on the dependent measure of concern. The third group consists of pupils who had a complete set of non-zero dependent measures (N=3,665). Lastly, the fifth group is our ANCOVA analytic sample which requires a complete covariate set and no less than five pupils per school in the sample (N=2,755).

The first group of 4,418 pupils may be considered as a "population" of longitudinally traced children who stayed in the same treatment group for the two year period. Although each restriction applied to this group reduces the size of the sample and the removal of zero score data increases the means, there should be no differential reduction for the various Sponsor FT/NFT groups.

Therefore we expect our analytic sample to be representative of the "population". If this is not the case we can then examine the remaining three groups of children in an attempt to ascertain which restriction is related to the lack of representativeness.

Tables II-9 through II-13 present the means and standard deviations required to examine changes in dependent measures as we reduce the sample from 4,418 to 2,755 children. The two heuristic devices discussed in Section 2.3.1 (difference between means of at least two standard errors and one quarter of a standard deviation) were used to make comparisons. As above, major emphasis will be placed on those Sponsor FT/NFT groups which satisfied both criteria. Those situations where the difference between means is greater than two standard errors but less than one quarter of a standard deviation will be indicated on Tables II-9 through II-13 and discussed if there appears to be a consistent pattern across outcome measures for a Sponsor FT/NFT group.

Overall, Far West Laboratory does not appear to have been greatly effected by our process of sample reduction for the ANCOVA analyses. However, for the three kindergarten tests explored (Spring PPVT, MAT Reading and Numbers) the FT group means appear to have been raised slightly. This is not true of the NFT group. No specific restriction appears to have caused the slight rise in test scores for the FT group. Although our FT analytic sample may have a slightly higher performance level than its parent "population", we should keep in mind that this sample's entry level on the WRAT was lower than its comparison NFT sample. While our sample restrictions may have helped to improve the appearance of results within this Sponsor's FT world, they did not result in the creation of comparable FT and NFT samples.

Although there are no cases in which the sample restrictions resulted in differences as great as one quarter of a standard deviation for the Kansas FT group, there does appear to be a consistent pattern across four of the five explored variables (kindergarten and first grade Reading and Math). It appears that the analytic sample's FT means are slightly higher than the FT "population" means on these tests. This phenomenon also occurs on one variable (first grade MAT Total Reading) for the Kansas NFT group. There is no one specific restriction which causes this trend to appear for all outcomes. For the kindergarten MAT reading test the exclusion of zero scores tends to raise FT scores. For the kindergarten MAT numbers test for FT and first grade

TABLE II-9

SPRING KINDERGARTEN PPVT MEANS, STANDARD DEVIATIONS AND SAMPLE SIZES BY SPONSOR BY FT/NFT FOR FIVE GROUPS OF CHILDREN

	Longitudinal Traced Sample with						All Dependent Measures Present			All Covariates Present			Analytic Sample		
	Zero Scores Included			Zero Scores Excluded											
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FT	51.592	9.960	[illegible]	51.611	9.914	2798	52.154	9.622	2377	52.288	9.352	1746	52.276	9.364	1724
FAR WEST	53.867	8.725	347	53.867	8.725	347	53.965	8.452	285	55.090[1]	7.863	223	55.005[1]	7.823	219
ARIZONA	53.722	9.913	414	53.722	9.913	414	54.148	9.494	385	52.669	9.781	260	52.669	9.781	260
BANK STREET	50.239	10.846	163	50.239	10.846	163	50.893	10.822	140	51.042	10.451	119	51.042	10.451	119
OREGON	50.968	8.617	155	51.299	7.593	154	51.472	7.488	127	51.432	7.539	95	51.432	7.539	95
KANSAS	48.579	9.393	522	48.579	9.393	522	49.162	9.160	445	48.843	9.191	324	48.831	9.218	320
HIGH SCOPE	52.125	8.556	263	52.125	8.556	263	51.789	8.878	185	52.185	8.903	168	52.185	8.903	168
FLORIDA	52.375	8.647	325	52.375	8.647	325	52.749	8.640	287	53.116	8.355	173	53.116	9.355	173
EDC	50.441	10.646	238	50.441	10.646	238	51.000	10.660	208	50.175	10.089	154	49.933	10.065	150
PITTSBURGH	59.033	7.084	182	59.033	7.084	182	59.465	6.637	170	59.562	6.639	153	59.562	6.639	153
SEDL	44.984	11.978	190	44.984	11.978	190	46.662	10.881	145	48.506[2]	8.254	77	48.343[2]	8.501	67
NFT	53.198	10.251	1572	53.198	10.251	1572	54.202	9.601	1288	54.320	9.338	1044	54.381	9.325	1031
FAR WEST	54.728	8.960	243	54.728	8.960	243	55.131	8.437	199	55.365	7.539	178	55.365	7.539	178
ARIZONA	58.197	8.898	269	58.197	8.898	269	58.727	8.389	249	57.828	9.162	169	57.807	9.211	166
BANK STREET	50.087	11.449	92	50.087	11.449	92	50.895	11.084	86	51.328	11.849	67	51.328	11.949	67
OREGON	51.190	9.734	105	51.190	9.734	105	51.088	10.192	91	51.279	10.171	86	51.256	10.406	92
KANSAS	50.523	9.416	149	50.523	9.416	149	51.394	8.641	132	51.646	8.625	113	51.646	8.625	113
HIGH SCOPE	52.315	9.228	203	52.315	9.228	203	52.768	8.874	168	52.862	8.930	145	52.862	8.930	145
FLORIDA	51.799	10.505	169	51.799	10.505	169	52.411[2]	10.125	112	54.321[2]	9.625	78	55.081[2]	9.103	74
EDC	51.731	10.595	160	51.731	10.595	160	54.134	9.537	119	53.344	9.799	96	53.344	9.799	96
PITTSBURGH	59.500	7.923	66	59.500	7.923	66	60.429	4.818	63	60.633	4.829	60	60.633	4.829	60
SEDL	48.147	11.710	116	48.147	11.710	116	49.623	10.910	69	52.615[2]	8.693	52	52.760[2]	8.754	50

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

TABLE II-10

SPRING KINDERGARTEN MAT READING SUBTEST MEANS, STANDARD DEVIATIONS, AND SAMPLE SIZES
BY SPONSOR BY FT/NFT FOR FIVE GROUPS OF CHILDREN

	Longitudinal Traced Sample with						All Dependent Measures Present			All Covariates Present			Analytic Sample		
	Zero Scores Included			Zero Scores Excluded											
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FT	14.201	5.674	2828	14.652	5.158	2741	14.837	5.132	2377	14.945	5.080	1746	14.943	5.090	1724
FAR WEST	13.462	5.302	351	13.897	4.790	340	14.133[1]	4.630	285	14.345[1]	4.744	223	14.311[1]	4.701	219
ARIZONA	14.524	5.582	416	14.701	5.379	411	14.945	5.343	385	14.527	5.254	260	14.527	5.254	260
BANK STREET	12.988	5.124	164	13.481	4.535	158	13.457	4.521	140	13.798[1]	4.167	119	13.798[1]	4.167	119
OREGON	17.385	6.429	156	17.610	6.154	154	17.394	6.475	127	17.863	6.478	95	17.863	6.478	95
KANSAS	15.621	5.698	528	16.204[1]	4.920	509	16.447[1]	4.844	445	16.309[1]	4.789	324	16.319[1]	4.809	320
HIGH SCOPE	13.018	5.334	271	13.622	4.640	259	13.676	4.627	185	13.839[1]	4.649	168	13.839[1]	4.649	168
FLORIDA	13.820	5.449	327	14.301	4.881	316	14.268[1]	4.885	287	14.584[1]	4.658	173	14.584[1]	4.658	173
EDC	13.315	4.796	241	13.597	4.431	236	13.668	4.516	208	13.675	4.627	154	13.567	4.591	150
PITTSBURGH	16.297	5.539	182	16.570	5.162	179	16.765	5.188	170	16.908	5.250	153	16.908	5.250	153
SEDL	10.839	5.265	192	11.626[1]	4.532	179	12.110[2]	4.343	145	12.390[2]	4.320	77	12.239[2]	4.499	67
NFT	13.491	5.797	1590	14.029	5.234	1529	14.422	5.182	1288	14.444	5.276	1044	14.446	5.290	1031
FAR WEST	13.844	4.994	243	13.959	4.852	241	14.025	4.986	199	13.972	5.067	178	13.972	5.067	178
ARIZONA	14.952	6.008	271	15.466	5.419	262	15.554	5.362	249	15.349	5.348	169	15.313	5.360	166
BANK STREET	16.409	5.207	93	16.587	4.942	92	16.733	5.051	86	16.672	5.223	67	16.672	5.223	67
OREGON	14.140	5.961	107	14.833	5.188	102	14.989	5.353	91	14.942	5.407	86	14.976	5.466	82
KANSAS	13.762	5.242	151	13.946	5.025	149	14.311	4.891	132	14.425	5.048	113	14.425	5.048	113
HIGH SCOPE	12.136	5.641	206	12.626	5.185	198	13.149[1]	4.963	168	13.117[1]	5.158	145	13.117[1]	5.158	145
FLORIDA	12.865	4.786	170	13.096	4.503	167	13.759[1]	4.315	112	14.372[2]	4.547	78	14.351[2]	4.665	74
EDC	12.876	5.627	161	13.374	5.117	155	13.588	5.091	119	13.677	5.338	96	13.677	5.338	96
PITTSBURGH	16.701	5.676	67	16.955	5.325	66	17.143	5.358	63	17.300	5.422	60	17.300	5.422	60
SEDL	8.579	5.826	121	10.701[2]	4.414	97	11.188[2]	4.163	69	11.404[2]	4.127	52	11.500[2]	4.161	50

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

TABLE II-11

SPRING KINDERGARTEN MAT NUMBERS SUBTEST MEANS, STANDARD DEVIATIONS, AND SAMPLE SIZES BY SPONSOR BY FT/NFT FOR FIVE GROUPS OF CHILDREN

	Longitudinal Traced Sample with						All Dependent Measures Present			All Covariates Present			Analytic Sample		
	Zero Scores Included			Zero Scores Excluded											
	X̄	S.D.	N	X̄	S.D.	N	X̄	S.D.	N	X̄	S.D.	N	X̄	S.D.	N
FT	12.847	6.667	2828	13.304	6.321	2731	13.561	6.341	2377	13.645	6.426	1746	13.654	6.439	1724
FAR WEST	11.641	6.129	351	12.457[1]	5.478	328	12.744[1]	5.520	285	12.848[1]	5.517	223	12.785[1]	5.487	219
ARIZONA	11.425	5.749	416	11.621	5.598	409	11.860	5.646	385	11.481	5.386	260	11.481	5.386	260
BANK STREET	10.128	5.041	164	10.786	4.465	154	10.757	4.131	140	10.714	4.289	119	10.714	4.289	119
OREGON	16.763	7.091	156	16.981	6.872	154	16.819	7.068	127	17.326	7.387	95	17.326	7.387	95
KANSAS	15.958	7.244	528	16.489	6.742	511	16.742[1]	6.741	445	16.750[1]	6.644	324	16.825[1]	6.648	320
HIGH SCOPE	10.601	5.093	271	11.050	4.697	260	11.276	4.878	185	11.381[1]	4.970	168	11.381[1]	4.970	168
FLORIDA	12.309	5.561	327	12.697	5.192	317	12.679	5.161	287	12.434	5.084	173	12.434	5.084	173
EDC	11.552	5.591	241	11.897	5.298	234	12.308	5.315	208	12.000	5.231	154	11.913	5.202	150
PITTSBURGH	17.863	8.355	182	18.162	8.094	179	18.453	7.966	170	18.634	7.972	153	18.634	7.972	153
SEDL	9.677	4.311	192	10.043	3.949	185	10.503[1]	3.760	145	11.221[2]	4.002	77	11.179[2]	3.861	67
NFT	12.462	6.530	1588	12.951	6.163	1528	13.392	6.254	1288	13.296	6.149	1044	13.332	6.169	1031
FAR WEST	14.218	7.059	243	14.578	6.769	237	14.658	6.845	199	14.489	6.714	178	14.489	6.714	178
ARIZONA	15.385	7.229	270	15.855	6.811	262	15.900	6.755	249	15.414	6.663	169	15.464	6.700	166
BANK STREET	11.677	4.879	93	11.804	4.749	92	11.895	4.775	86	11.940	4.386	67	11.940	4.386	67
OREGON	12.364	6.949	107	12.845	6.630	103	13.066	6.623	91	12.919	6.516	86	13.024	6.615	82
KANSAS	11.722	6.146	151	11.959	5.974	148	12.447	6.040	132	12.124	5.614	113	12.124	5.614	113
HIGH SCOPE	9.976	5.610	206	10.816	5.000	190	11.143[1]	5.120	168	11.372[2]	5.349	145	11.372[2]	5.349	145
FLORIDA	12.076	6.251	170	12.673	5.780	162	13.723[2]	6.118	112	14.923[2]	6.316	78	15.203[2]	6.350	74
EDC	11.453	5.371	161	11.597	5.247	159	12.445	5.489	119	12.281	5.719	96	12.281	5.719	96
PITTSBURGH	13.478	6.354	67	14.109	5.766	64	14.206	5.759	63	14.400	5.788	60	14.400	5.788	60
SEDL	9.550	4.587	120	10.324[1]	3.833	111	10.623	4.177	69	10.769[2]	3.874	52	10.880[2]	3.880	50

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

TABLE II-12

SPRING FIRST GRADE MAT TOTAL READING SUBTEST MEANS, STANDARD DEVIATIONS, AND SAMPLE SIZES BY SPONSOR BY FT/NFT FOR FIVE GROUPS OF CHILDREN

	Longitudinal Traced Sample with Zero Scores Included			Longitudinal Traced Sample with Zero Scores Excluded			All Dependent Measures Present			All Covariates Present			Analytic Sample		
	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FT	40.911	17.470	2828	41.408	16.979	2794	42.266	16.894	2377	42.651	17.219	1746	42.560	17.238	1724
FAR WEST	36.675	16.067	351	36.885	15.870	349	37.547	15.636	285	37.803	15.721	223	37.785	15.737	219
ARIZONA	41.397	17.470	416	41.799	17.069	412	42.309	17.158	385	41.523	16.942	260	41.523	16.942	260
BANK STREET	38.604	16.588	164	39.323	15.870	161	38.679	15.177	140	39.277	15.618	119	39.277	15.618	119
OREGON	53.506	17.111	156	53.852	16.613	155	52.717	16.786	127	53.600	17.447	95	53.600	17.417	95
KANSAS	43.953	17.803	528	44.373	17.358	523	45.627[2]	17.033	445	46.664[2]	17.125	324	46.537[2]	17.168	320
HIGH SCOPE	32.229	13.854	271	32.835	13.251	266	33.649	13.043	185	34.196	13.257	168	34.196	13.257	168
FLORIDA	39.232	15.564	327	39.718	15.030	323	39.965	15.053	287	39.855	15.517	173	39.855	15.517	173
EDC	40.091	15.829	241	40.258	15.647	240	40.538	16.001	208	40.825	16.133	154	40.393	15.952	150
PITTSBURGH	52.923	19.937	182	53.511	19.243	180	54.818	18.644	170	55.255	18.861	153	55.255	18.861	153
SEDL	35.724	13.692	192	37.076	12.008	185	38.738[1]	11.759	145	38.649[2]	10.548	77	37.179	9.550	67
NFT	42.523	18.512	1590	43.230	17.828	1564	45.233	17.914	1288	45.377	17.855	1044	45.454	17.877	1031
FAR WEST	41.815	17.785	243	42.873	16.696	237	43.834	17.121	199	43.983	17.305	178	43.983	17.305	178
ARIZONA	54.399	17.653	271	54.399	17.653	271	55.622	17.227	249	54.272	17.135	169	54.205	17.137	166
BANK STREET	47.323	17.884	93	47.837	17.277	92	48.477	17.232	86	48.224	17.270	67	48.224	17.270	67
OREGON	45.888	19.275	107	46.762	18.368	105	47.582	18.512	91	47.419	18.555	86	48.000	18.741	82
KANSAS	37.993	17.593	151	39.840	15.830	144	40.848[1]	15.697	132	41.646[1]	16.233	113	41.646[1]	16.233	113
HIGH SCOPE	37.893	15.529	206	38.265	15.141	204	39.131	15.227	168	39.621	15.532	145	39.621	15.532	145
FLORIDA	36.053	17.768	170	36.922	17.062	166	40.848[2]	17.225	112	45.154[2]	18.137	79	45.568[2]	18.496	74
EDC	38.112	16.271	161	38.591	15.795	159	40.782	16.031	119	41.479	16.498	96	41.479	16.498	96
PITTSBURGH	57.955	15.693	67	57.955	15.693	67	58.317	16.053	63	58.867	16.199	60	58.867	16.199	60
SEDL	30.628	11.070	121	31.143	10.414	119	30.725	10.696	69	30.327	10.041	52	30.640	10.111	50

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

TABLE II-13

SPRING FIRST GRADE MAT MATHEMATICS SUBTEST MEANS, STANDARD DEVIATIONS, AND SAMPLE SIZES BY SPONSOR BY FT/NFT FOR FIVE GROUPS OF CHILDREN

	Longitudinal Traced Sample with						All Dependent Measures Present			All Covariates Present			Analytic Sample		
	Zero Scores Included			Zero Scores Excluded											
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N
FT	30.883	13.358	2828	31.203	13.050	2799	31.894	12.806	2377	32.030	13.181	1746	31.958	13.138	1724
FAR WEST	29.399	11.680	351	29.652	11.404	348	29.937	10.858	285	29.785	11.257	223	29.800	11.250	219
ARIZONA	28.897	12.094	416	28.966	12.025	415	29.582	12.005	385	29.019	12.116	260	29.019	12.116	260
BANK STREET	28.067	12.633	164	28.414	12.316	163	28.300	11.556	140	27.849	11.595	119	27.849	11.595	119
OREGON	32.096	12.018	156	32.303	11.774	155	31.575	11.809	127	31.526	12.317	95	31.526	12.317	95
KANSAS	34.466	13.175	528	34.597	13.026	526	35.344	12.795	445	36.657[1]	13.268	324	36.688[1]	13.336	320
HIGH SCOPE	21.188	10.788	271	21.833	10.287	263	23.081[1]	10.282	185	23.500[1]	10.366	168	23.500[1]	10.366	168
FLORIDA	30.697	11.392	327	30.981	11.042	324	30.951	11.032	287	31.312	11.253	173	31.312	11.253	173
EDC	34.714	13.939	241	35.449	13.125	236	35.553	12.996	208	35.597	12.889	154	35.267	12.871	150
PITTSBURGH	39.956	14.734	182	40.177	14.470	181	41.359	13.801	170	41.471	14.010	153	41.471	14.010	153
SEDL	30.063	14.776	192	30.540	14.393	189	31.800	13.872	145	30.636	13.788	77	29.179	12.436	67
NFT	30.861	13.636	1590	31.414	13.111	1562	32.708	12.908	1288	32.459	12.981	1044	32.467	13.016	1031
FAR WEST	32.778	12.442	243	33.050	12.127	241	33.211	12.089	199	32.674	12.166	178	32.674	12.166	178
ARIZONA	37.959	12.088	271	37.959	12.088	271	38.755	11.799	249	37.769	11.795	169	37.717	11.889	166
BANK STREET	30.355	11.810	93	30.355	11.810	93	30.895	11.471	86	30.970	11.396	67	30.970	11.396	67
OREGON	34.523	12.795	107	38.849	12.402	106	35.473	12.462	91	35.628	12.420	86	36.146	12.394	82
KANSAS	26.748	13.087	151	26.748	13.087	151	28.174	12.698	132	28.416	12.749	113	28.416	12.749	113
HIGH SCOPE	27.102	12.526	206	27.776	11.916	201	28.494	11.951	168	28.731	12.339	145	28.731	12.339	145
FLORIDA	27.094	15.292	170	30.105[1]	12.991	153	31.446[2]	13.449	112	32.859[2]	14.175	78	32.608[2]	14.437	74
EDC	28.050	13.217	161	28.225	13.070	160	30.555	12.828	119	29.896	13.289	96	29.896	13.289	96
PITTSBURGH	41.149	11.927	67	41.149	11.927	67	41.016	12.150	63	40.917	12.431	60	40.917	12.431	60
SEDL	23.132	10.870	121	23.521	10.532	119	25.159	10.393	69	24.692	10.186	52	24.840	10.267	50

[1] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors.

[2] Difference between means of the random sample and means of those present at three testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

Reading for FT and NFT, scores are increased by requiring complete sets of non-zero dependent measures. First grade Math scores are slightly higher for FT pupils who have a complete covariate set present.

High Scope's FT and NFT ANCOVA analytic samples do not seem to be at great deviance from their respective longitudinal traced "populations". On two of the kindergarten tests (Reading and Numbers) the FT group mean may have been raised slightly by our requirement that a complete set of covariates be present. For these same two tests the NFT means seem to have been raised also, apparently as a consequence of zero score exclusion of the Numbers subtest, and the requirement that all dependent measures be non-zero on the Reading test. For the first grade Mathematics test, the FT sample's mean also was raised slightly by this later restriction. Thus, while there appears to be no great deviance introduced by our sample restrictions for High Scope, there is a pattern which suggests that both FT and NFT means may have been raised slightly.

The University of Florida's NFT ANCOVA analytic sample means have been consistently raised while the FT means have not. The NFT phenomenon appears to be related to the restriction that a complete set of non-zero dependent measures be present. Under these circumstances it is impossible to demonstrate the true FT/NFT effect since we have differentially removed the lower performers from their NFT comparison sample. This becomes crucial in that Florida appears to be affected in both reading and math at both the kindergarten and first grade levels. It is possible that the Florida FT effect may be unfairly conservative.

Both FT and NFT group means for SEDL have been raised on the three kindergarten tests by our sample restrictions. The FT mean on the first grade Total Reading was also raised. There appears to be no specific restriction which resulted in this occurrence.

The Sponsors not discussed in this section (Arizona, Bank Street, Oregon, EDC and Pittsburgh) show no evidence of changes in the FT/NFT samples which are related to the selection criteria examined here.

The restrictions employed in the sample definition process have resulted in differential attrition for the Sponsor FT/NFT groups. The fact that differential attrition has occurred at the Sponsor level (some affected, others not), at the treatment level within Sponsor (in some Sponsors both the FT and the NFT; in others only the FT or NFT), and at the test level (some tests for some groups, other tests for other groups) severely limits our ability to both generalize to larger FT/NFT groups within Sponsor and to compare Sponsor results. Such findings become crucial in a situation such as the University of Florida's where at this final sample definition stage we appear to have consistently increased the NFT means. For the University of Kansas complications exist not only at this last stage but also at the first stage of requiring three testing points.

2.3.3 Comparison of 1974 Analytic Sample with Fall 1971 Tested Population and 1973 Analytic Sample

Up to this point we have been studying the specific steps in the sample definition process, ascertaining whether or not each of the several steps resulted in a bias or nonrepresentativeness. We now turn to a direct comparison of the current ANCOVA analytic sample with the population of Cohort III kindergarteners tested in the Fall of 1971. Using the Fall WRAT (Reading and Spelling subtests)[4] and PPVT we will assess the extent to which our current analytic sample is representative of the larger population on entering achievement. We will also utilize mother's education and family income to assess the comparability of the current samples and the population on SES. Simultaneous with this comparison we will also include data on the 1973 analytic sample. Focusing on the same entry level and SES measures mentioned above, we will examine the comparability of this year's sample to last year's sample. By examining these three groups of children (Fall 1971 tested population, last year's analytic sample and this year's analytic sample) we will be able to note instances of differential attrition within and across Sponsors.

The relevant data are presented on Tables II-14 through II-23. When comparing the analytic sample to the population, we have once again employed the two heuristic devices of examining whether the difference in means of any two groups is greater than two standard errors or greater than one quarter of a standard deviation. For a comparison of the two samples, t tests were used as a heuristic device when the sample means were at least a quarter of a standard deviation apart. Note carefully that these t values are only indicators of possible significant differences between means; when the standard deviations are unequal, the statistical significance of it is unknown.

---

[4] We made a slight change from last year to this year in our scoring specifications for the WRAT arithmetic subtest in order to duplicate the scoring used by the test publisher. Therefore we cannot compare the samples using the WRAT arithmetic subtest or the total WRAT score (which includes this subtest score).

Far West. For Far West's NFT group there appears to be a consistent pattern across all three entry tests: the current NFT sample means are higher than the NFT population means. This is not true of the current FT sample. As a result, although Far West's entering population had comparable FT and NFT groups, as measured by pretests, the NFT group of the current sample seems to have slightly higher pretest scores than the FT group.

There are no differences between the FT/NFT sample and population for this Sponsor on mothers' education. On income, while there are no significant differences, the FT sample mean appears to be somewhat lower and the NFT sample mean somewhat higher than the corresponding population means. Thus, income differences in the Far West program, like pretest differences, increase from population to sample.

The current analytic sample also differs from last year's analytic sample on pretest scores. This year's FT group has a lower mean entry level on the PPVT and WRAT reading subtest than last year's FT group. On the other hand, the current NFT group has a higher mean entry level on the WRAT spelling subtest than last year's NFT group. The net effect of these sample changes is to increase the discrepancy between the FT/NFT groups on pretest measures. There are no corresponding differences, however, on either of the SES indices.

Arizona. The current FT/NFT sample for the University of Arizona seems to be fairly representative of the Arizona FT/NFT population which entered kindergarten in the Fall of 1971. For both the entering population and this year's analytic sample there is a significant FT disadvantage on all of the pretest and SES measures.

Comparing this year's and last year's analytic samples, there are no

TABLE II-14

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR FAR WEST BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	43.91	12.37	871	49.33[3]	7.98	316	46.52[1]	11.10	219
	NFT	43.42	13.49	894	49.84	8.12	313	48.42[1]	9.85	178
FALL KINDERGARTEN WRAT READING SUBTEST	FT	13.96	5.79	864	15.20[3]	5.83	316	13.94	5.29	217
	NFT	13.74	5.87	889	14.62	6.05	313	14.68	6.01	178
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	5.31	3.53	715	4.75	3.90	316	5.33	3.58	184
	NFT	6.01	4.01	769	5.70[3]	4.24	313	6.80[1]	3.87	164
MOTHER'S EDUCATION INDEX	FT	5.52	1.01	572	5.68[3]	0.93	302	5.40	0.97	215
	NFT	5.53	1.03	637	5.63	0.93	303	5.53	1.19	174
FAMILY INCOME INDEX	FT	8.51	6.65	575	9.37[3]	6.74	316	8.13	6.31	209
	NFT	11.00	7.65	649	11.61	7.51	313	11.89	7.44	175

[1] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors.

[2] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

TABLE II-15

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR ARIZONA BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N
FALL PPVT	FT	43.91	13.78	918	48.56[3]	8.53	334	43.43	12.45	260
	NFT	47.77	13.15	837	52.19[3]	8.78	275	50.03[1]	12.59	166
FALL KINDERGARTEN WRAT READING SUBTEST	FT	14.22	6.37	897	14.56	6.65	334	14.65	6.55	256
	NFT	15.74	6.75	829	15.94	6.64	275	16.37	6.62	166
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	5.96	3.85	804	5.47	4.07	334	5.70	3.71	234
	NFT	7.01	4.30	770	6.78	4.53	275	7.47	4.67	155
MOTHER'S EDUCATION INDEX	FT	5.46	1.19	555	5.73	1.05	317	5.51	1.06	248
	NFT	5.90	1.04	540	6.01	1.05	265	6.01	1.03	162
FAMILY INCOME INDEX	FT	10.24	7.03	566	10.90	7.18	334	10.66	6.77	257
	NFT	14.23	7.26	546	14.91	7.32	275	15.35	7.17	158

[1] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

measurable differences on the WRAT subtests, income or mother's education for FT and NFT groups. However, on the entering PPVT test, last year's sample (both FT and NFT groups) had higher scores than this year's sample. Despite sample difference, the FT disadvantage remains constant.

Bank Street. Current FT/NFT samples for Bank Street are fairly representative of the Fall 1971 tested populations on all three pretests. On the other hand, there is a substantial difference between the NFT sample and population on mother's education, with the NFT sample being much lower on this variable.

It would appear that for some reason, there is a great deal of attrition in the NFT group on the part of children with highly educated mothers. Either this group is fairly mobile, children from this group are more apt to be moved into the FT program, or for some administrative reasons they fail to be tested. In any event, this dramatic shift suggests that the sample is not fully representative, despite apparent comparability with the population on pretests.

Last year's sample is markedly different from the current sample on two of the pretests. The 1973 NFT group had higher scores than the 1974 NFT group on the Fall WRAT reading subtest. On the PPVT, both FT and NFT groups for last year had higher scores than the groups currently in use.

Consistent with these pretest differences, the FT/NFT groups for 1973 were higher on level of mother's education than this year's groups. Thus, the sample selection procedure yielded lower achieving, lower SES children for the present analyses, especially for the NFT group.

Oregon. Oregon's current FT/NFT analytic sample is representative of the population tested in the Fall of 1971 on both pretest and SES measures.

On SES measures, this year's sample is also similar to last year's sample. In both cases, the FT group is lower than the NFT group on mean income and mother's education. The most important difference between the current Oregon sample and that of last year is on the PPVT. Both FT and NFT groups had considerably higher scores last year on this measure. Despite this difference between samples, the Oregon pattern of FT/NFT comparability on entry level remains constant over time.

TABLE II-16

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR BANK STREET BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	40.53	13.70	756	47.11^{3}	7.86	224	43.34^{1}	11.43	119
	NFT	39.97	14.74	764	50.00^{3}	8.08	135	41.34	15.51	67
FALL KINDERGARTEN WRAT READING SUBTEST	FT	13.98	6.43	744	14.88	6.36	224	14.18	6.18	116
	NFT	14.31	6.29	752	15.59^{3}	6.07	135	13.00^{1}	5.22	67
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	5.28	3.58	646	5.22	3.82	224	5.21	3.56	103
	NFT	5.42	3.65	677	6.09	3.66	135	5.84	3.46	62
MOTHER'S EDUCATION INDEX	FT	5.48	1.05	432	5.51	0.98	212	5.26^{1}	0.97	108
	NFT	5.36	1.28	327	5.60^{3}	1.19	131	4.58^{2}	1.73	66
FAMILY INCOME INDEX	FT	11.09	7.08	436	10.85	6.72	224	10.31	6.39	114
	NFT	13.27	7.03	325	14.62	6.57	135	13.54	6.69	63

[1] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors.

[2] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

TABLE II-17

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR OREGON BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	42.40	11.63	775	45.66[3]	7.48	162	42.50	10.49	95
	NFT	40.12	14.31	665	46.11[3]	8.47	124	41.61	12.43	82
FALL KINDERGARTEN WRAT READING SUBTEST	FT	14.20	6.22	770	14.52	5.55	162	13.81	5.46	94
	NFT	14.72	6.46	654	15.41	6.71	124	14.59	6.47	82
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	6.15	3.79	705	5.88	3.90	162	6.21	3.59	90
	NFT	5.80	3.93	579	4.74	3.71	124	5.33	3.54	67
MOTHER'S EDUCATION INDEX	FT	5.48	1.07	406	5.44	1.03	153	5.33	1.03	91
	NFT	5.65	1.13	375	5.77	1.01	119	5.68	1.09	78
FAMILY INCOME INDEX	FT	8.13	6.44	408	6.19	4.89	162	7.14	5.84	93
	NFT	12.49	7.28	363	12.62	7.43	124	12.36	7.35	77

---

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

Kansas. As measured by the Fall WRAT subtest scores, the current FT/NFT sample for Kansas resembles the Fall 1971 tested population. On the PPVT, however, this year's FT and NFT groups (especially the NFT group) have higher PPVT scores than the original groups they represent. The current NFT group also has a somewhat higher level of mother's education than the original tested population. Thus, the current Kansas sample appears to be somewhat less disadvantaged than the original population, especially the NFT group.

This year's sample does not differ greatly from last year's sample, except on the PPVT. Last year's FT/NFT groups were even higher than this year's on this pretest measure and, therefore, even less representative of the entering populations. While differences do exist between analytic samples, they do not affect the FT/NFT mismatch within sample and will interfere only with the ability to generalize FT/NFT contrasts for Kansas.

High Scope. In the current analytic sample, High Scope seems to have FT and NFT groups which are fairly representative of the 1971 entering population on all five of the measures examined.

There are some marked differences between the 1973 and 1974 samples for High Scope. On the PPVT, the current FT/NFT groups have lower scores than those of last year. This trend is also found when we compare the samples on mother's education: this year's sample of mothers is less educated than last year's for both FT and NFT. On the WRAT spelling subtest, however, this trend is reversed; the current sample has higher scores than last year's sample. In sum, High Scope is represented differently in this year's analysis than it was in last year's analysis. However, the within Sponsor FT/NFT comparisons seem to be similar across years, with the exception of the WRAT reading subtest where the current sample shows a greater FT disadvantage than was found last year.

University of Florida. Comparing the University of Florida's Fall 1971 tested population with the current analytic sample, the former is significantly lower on the PPVT. Moreover, the current NFT sample has higher scores on the WRAT spelling subtest than the NFT population. There are no other differences between the population and sample on either pretests or SES measures.

TABLE II-18

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR KANSAS BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	36.84	12.76	1204	44.34[3]	8.29	337	39.53[1]	11.53	320
	NFT	34.71	14.25	1036	43.96[3]	9.47	125	40.32[1]	12.39	113
FALL KINDERGARTEN WRAT READING SUBTEST	FT	13.13	5.81	1192	14.17	5.82	337	13.58	5.46	318
	NFT	12.34	5.86	991	11.90	5.36	125	11.60	5.36	113
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	5.48	3.10	1029	5.58	4.04	337	5.68	3.83	284
	NFT	4.90	3.76	817	4.53	4.17	125	5.06	3.99	100
MOTHER'S EDUCATION INDEX	FT	5.37	1.05	673	5.45	1.01	306	5.31	1.07	287
	NFT	5.00	1.21	427	5.36	.97	110	5.27[1]	1.04	100
FAMILY INCOME INDEX	FT	8.26	6.73	706	8.26	6.72	337	7.77	6.53	308
	NFT	9.45	7.53	444	9.04	7.31	125	8.29	7.17	110

---

[1] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

TABLE II-19

MEANS AND STANDARD DEVIATIONS ON
PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR HIGH SCOPE BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	39.67	13.26	532	45.80[3]	8.49	185	42.18[1]	11.82	168
	NFT	41.81	13.58	580	47.88[3]	8.16	150	41.92	13.22	145
FALL KINDERGARTEN WRAT READING SUBTEST	FT	13.01	5.83	531	13.90	5.69	185	12.93	5.43	167
	NFT	14.75	6.66	575	14.27	6.43	150	14.30	6.31	145
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	4.93	3.44	457	4.60	3.61	185	5.05	3.52	147
	NFT	5.63	3.88	510	4.63[3]	3.75	150	5.67	3.65	125
MOTHER'S EDUCATION INDEX	FT	5.41	1.1[illegible]	410	5.73[3]	.90	173	5.28	1.23	157
	NFT	5.37	1.21	355	5.60[3]	1.06	144	5.28	1.26	139
FAMILY INCOME INDEX	FT	7.55	6.36	431	7.98	6.86	185	7.93	6.59	167
	NFT	10.92	7.30	368	11.22	7.23	150	11.32	7.15	142

---

[1] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors.

[3] A t test indicates a possible significant difference betwsen the 1973 and 1974 sample means.

TABLE II-20

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR FLORIDA BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N
FALL PPVT	FT	40.83	13.39	679	48.53[3]	8.43	255	44.86[2]	11.87	173
	NFT	37.45	15.03	603	48.39	9.64	104	46.84[2]	13.05	74
FALL KINDERGARTEN WRAT READING SUBTEST	FT	12.90	5.28	658	13.08	6.32	255	13.66	6.11	169
	NFT	11.56	5.48	580	13.13	6.03	104	12.45	6.02	73
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	5.21	3.56	550	4.91	4.26	255	5.22	3.60	143
	NFT	5.18	3.91	483	6.34	4.71	104	6.94[2]	4.42	68
MOTHER'S EDUCATION INDEX	FT	5.27	1.06	439	5.48	.88	243	5.30	1.03	163
	NFT	5.26	1.10	316	5.68[3]	.95	101	5.32	1.01	71
FAMILY INCOME INDEX	FT	9.15	7.31	439	10.15[3]	7.59	255	8.36	7.05	165
	NFT	11.81	7.48	309	13.21	7.93	104	13.10	7.35	67

[2] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

The 1973-74 analytic samples differ in a number of ways. The 1973 FT sample for Florida is higher than the 1974 FT sample on both the PPVT and income level. On the other hand, the 1973 NFT sample is higher than the 1974 NFT sample on mother's education. The net effects of these sample differences are that this year's FT sample is slightly lower than last year's on both the PPVT and income indices and that concomitantly there is a greater discrepancy between this year's FT/NFT groups on these variables.

EDC. EDC's current analytic sample is fairly comparable to the entering population of EDC pupils. The one instance of dissimilarity is the NFT mean PPVT score which is higher in the current sample than it is for the population. This difference creates significant FT disadvantage on the PPVT in the current sample which does not exist in the entering population and is not found on the other pretest or SES measures.

The only significant difference between the current analytic sample and last year's sample is on the PPVT. Both FT and NFT groups have higher PPVT scores in last year's sample than in this year's sample.

A nonsignificant but consistent pattern of sample to sample differences is found on the WRAT reading subtest and the family income index. On these two measures the current NFT group has slightly lower scores than last year's NFT group. This results in a closer FT/NFT match in the current analysis than was found last year.

Pittsburgh. It was noted above in Section 3.1 that a traced longitudinal sample for Pittsburgh has higher entry scores than a cross sectional sample. This is reflected in the current FT/NFT analytic sample for the University of Pittsburgh which has markedly higher pretest scores than the entering population. So, too, the FT sample is higher than the population on both mother's education and income.

A comparison of the 1973 and 1974 analytic samples reveals a pattern which has not been found in any of the other Sponsors. The current FT/NFT groups have consistently higher scores than last year's FT/NFT groups. On all pretest measures, moreover, the current FT group is higher on both measures of SES than last year's FT group. Lower achieving children drop out of the program or fail to be tested at a greater rate than the higher achieving children. This is not a universal process across all Sponsors.

TABLE II-21

MEANS AND STANDARD DEVIATIONS ON
PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN
FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR
EDC BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N
FALL PPVT	FT	41.10	13.44	681	46.66[3]	8.22	253	41.30	12.29	150
	NFT	41.17	14.01	647	50.49[3]	8.97	148	46.35[2]	12.37	96
FALL KINDERGARTEN WRAT READING SUBTEST	FT	13.65	6.34	664	14.33	6.36	253	14.28	6.10	148
	NFT	14.65	6.59	633	15.64	6.75	148	14.73	6.51	94
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	5.35	3.87	585	4.93	3.87	253	5.74	4.03	136
	NFT	5.48	3.77	549	5.35	4.14	148	5.52	3.80	83
MOTHER'S EDUCATION INDEX	FT	5.52	.94	560	5.63	.92	250	5.60	.89	145
	NFT	5.59	.98	337	5.75	1.02	147	5.55	1.18	93
FAMILY INCOME INDEX	FT	10.93	6.81	555	11.76	6.72	253	11.59	7.02	142
	NFT	12.83	6.95	340	14.04	6.74	148	12.72	7.01	95

---

[2] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

TABLE II-22

MEANS AND STANDARD DEVIATIONS ON
PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR PITTSBURGH BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	47.08	10.84	396	51.10[3]	8.45	268	53.75[2]	7.95	153
	NFT	47.42	12.28	499	52.82	7.22	164	54.48[2]	6.24	60
FALL KINDERGARTEN WRAT READING SUBTEST	FT	13.83	5.52	394	14.41	5.69	268	15.45[2]	6.01	153
	NFT	13.83	6.44	493	13.95[3]	6.26	164	16.71[2]	7.13	59
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	6.45	4.44	355	6.74	4.70	268	8.45[2]	4.70	148
	NFT	6.73	4.26	450	6.90	4.12	164	7.47	3.43	58
MOTHER'S EDUCATION INDEX	FT	5.74	.84	300	5.78	.81	258	5.90[1]	.88	152
	NFT	5.95	.95	366	6.04	.95	160	5.97	.91	59
FAMILY INCOME INDEX	FT	10.24	7.25	303	10.62[3]	7.36	268	13.50[2]	7.02	145
	NFT	14.20	6.52	368	14.46	6.19	164	14.78	6.19	60

[1] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors.

[2] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

We are analyzing a somewhat distorted sample of Pittsburgh children.

SEDL. The last Sponsor, SEDL, seems to have a sample for this year's analysis which is not representative of the population of students who entered the program in the Fall of 1971 on pretest measures. The current FT group has a lower mean score on the two WRAT subtests than its respective FT population; the current NFT group has a higher mean PPVT score than its respective NFT population. In Section 3.1 above this was shown to be related to the restriction of longitudinal data. The result is an FT disadvantage on all three pretest measures which is not found in the entering population.

On SES measures, the current sample more closely resembles the entering population. There is a slight but nonsignificant trend for the sample to have higher SES indicators than the parent population. It is interesting that the sample to population pretest differentials, which result in a nonrepresentative analytic sample, are not found in family income or mother's education.

The current FT sample for SEDL has lower PPVT and WRAT reading scores than the previous sample; the current NFT sample has higher WRAT spelling scores than last year's sample. The slight sample to sample variations on the two SES measures are not significant. In short SEDL's current sample, which is the smallest sample in the analysis, is not similar to either the entering population or last year's sample especially on pretest measures.

Summary. In summary, it appears that there are differences among and within Sponsors on the representativeness of the FT/NFT groups in the current sample. For three of the ten Sponsors (Arizona, Oregon, and High Scope), the current analytic sample is comparable to the Fall 1971 tested population on both entry level and SES measures. While Bank Street is representative in terms of entry level, it is not on mother's education, at least for the NFT group. For three other Sponsors, (Kansas, Florida, and Pittsburgh) the FT/NFT groups in the current sample are not representative; for these Sponsors, the FT/NFT groups in the current sample show higher pretest scores than the Fall 1971 tested populations. Finally, for Far West, EDC, and SEDL (and Florida on WRAT spelling), the FT/NFT groups in the current sample

TABLE II-23

MEANS AND STANDARD DEVIATIONS ON PRETEST AND SES MEASURES FOR THE POPULATION OF CHILDREN TESTED IN FALL 1971 AND THE 1973 AND 1974 ANALYTIC SAMPLE FOR SEDL BY FT/NFT

		Fall 1971 Tested Population			1973 Analytic Sample			1974 Analytic Sample		
		$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.	N
FALL PPVT	FT	32.80	17.13	983	44.51^{3}	7.88	76	34.75	13.81	67
	NFT	33.68	16.14	667	46.08	8.85	61	44.20^{2}	11.03	50
FALL KINDERGARTEN WRAT READING SUBTEST	FT	12.19	5.66	910	12.16^{3}	4.90	76	9.97^{2}	4.72	66
	NFT	12.09	5.35	635	11.33	3.89	61	11.25	3.97	49
FALL KINDERGARTEN WRAT SPELLING SUBTEST	FT	6.44	4.25	797	3.91	3.27	76	4.77^{2}	3.72	52
	NFT	5.68	3.77	551	3.80^{3}	3.56	61	6.03	3.42	40
MOTHER'S EDUCATION INDEX	FT	4.45	1.56	417	5.23	1.40	75	4.77	1.40	65
	NFT	4.69	1.41	423	5.08	1.31	61	4.76	1.45	49
FAMILY INCOME INDEX	FT	8.06	6.41	418	9.32	7.41	76	8.27	6.95	67
	NFT	9.41	6.64	427	10.87	7.21	61	10.26	7.16	50

[2] Difference between means of the Fall 1971 tested population and the means of those present at all testing times equals or exceeds two standard errors and one-quarter of a standard deviation.

[3] A t test indicates a possible significant difference between the 1973 and 1974 sample means.

are differentially representative. In these cases, the current NFT sample is higher on pretests than the NFT population. Moreover, the FT sample for SEDL is lower than the FT population on pretest scores. For these Sponsors, as well as Bank Street, the effects of attrition on the current sample are not only to limit the generalizability of the findings but to increase the difficulty in making appropriate FT/NFT comparisons.

The current samples for several Sponsors differ not only from the 1971 Fall tested populations but also from last year's analytic samples as well. The discrepancy between the 1973 and 1974 samples is no doubt associated with the different selection criteria employed for each year's analysis. Unlike this year, last year required complete data on family income and mother's education, permitted zero test scores, was limited to Black and White students, and repuired only cross sectional (one year) data. These criteria probably alter the sample characteristics of different Sponsors in different ways.

In general, last year's analytic sample was higher in entering achievement than this year's, at least as measured by the Fall kindergarten PPVT. The FT/NFT groups for five of the ten Sponsors (Arizona, Bank Street, Oregon, Kansas, and EDC) all showed higher scores last year than this year on this pretest measure. Only one Sponsor, Pittsburgh, showed lower scores last than this. For the four other Sponsors (Far West, High Scope, Florida, and SEDL) last year's sample and this year's sample differ in rather complicated ways.

2.4 DISCUSSION

This Section in addition to defining the current analytic sample, has addressed one of two issues central to our goal of estimating Sponsor effects.[1] The issue of generalizability--how well does our analytic sample represent the Fall 1971 tested population with which a Sponsor is working. To the extent that our sample is not representative of this FT population, our estimates of Sponsor effectiveness are not generalizable. The second issue which is addressed in Appendix A of Volume IIA is that of FT/NFT mismatch. The greater the FT/NFT comparisons within any one Sponsor, the more difficult it is to make appropriate FT/NFT comparisons for that Sponsor in the ANCOVA design.

---

[1] For a full description of the analytic sample see Appendix A Volume IIA.

The one issue which has not been addressed directly is the feasibility of drawing Sponsor-to-Sponsor comparisons. The reader will note, however, that, in effect, the question of cross-Sponsor comparisons has already been answered in the course of this Section and Appendix A. The Sponsors' samples differ from each other in all possible ways. Sponsors vary in the extent to which the sample represents the population as well as the magnitude, direction, and type of FT/NFT mismatch. The Sponsors are working with different children from different families in different locations under different circumstances. Since the meaning of any Sponsor effect is inextricably tied to unique characteristics of the sample on which it is based, we cannot compare effects across Sponsors.

CHAPTER III

VARIABLES

3.0 INTRODUCTION

The kindergarten and two year analyses used variables in three different ways. Treatment, i.e., FT/NFT membership, serves as the single predictor variable. An achievement test and selected socioeconomic measures were used as covariates. Three achievement tests, two affective tests and number of days absent served as outcome measured at the end of kindergarten. The first grade outcome measures included one achievement test and number of days absent in first grade. Subtest rather than total scores were used for some of the outcome measures at both the kindergarten and first grade levels. Both the covariates and outcome measures are discussed in greater detail below.

3.1 COVARIATES

The covariate selection process which culminated in the final selection of the limited set of covariates which were used in the various analyses is discussed in Appendix C. For the sponsor level kindergarten study, the total score on the fall WRAT served as the academic achievement covariate. The content characteristics of this test are discussed below under Outcome Measures. Two socioeconomic variables were used as covariates in this kindergarten study. They are Mother's Education and Family Income. For the site level kindergarten study, the achievement covariate was once again the fall WRAT total score, Mother's Education and an associated "missingness" code which indicates presence or absence of the datum on that variable which the socioeconomic covariates used. For the two year study which explored the first grade outcomes at the sponsor level, the kindergarten level fall WRAT was the academic achievement covariate. Because of heterogeneity problems, no socioeconomic covariates were used. At the site level, the fall WRAT was the academic achievement covariate. Mother's education and its associated "missingness" code once again served as the socioeconomic covariates.

3.2 OUTCOME MEASURES

3.2.1 Tests Administered by SRI

In the Spring in 1972, as the Cohort III pupils were completing kindergarten, three tests were administered in the achievement domain and two in the affective. The achievement tests[1] included:

PPVT, the Peabody Picture Vocabulary Test, Form A;

MAT, the Metropolitan Achievement Test, Primer, Form F; and

WRAT, a shortened version of the Wide Range Achievement Test.

The affective tests were:

Gumpgookies, a shortened, sixty-item version; and

Locus of Control.

In the Spring of 1973, as these pupils were completing the first grade, they were given:

MAT, the Metropolitan Achievement Test, Primary I, Form F.

Two additional variables served as outcome measures in these analyses:

the number of days absent in kindergarten; and

the number of days absent in first grade.

3.2.2 Domains Examined in the Achievement Batteries

As the subtest names suggest,[2] the achievement batteries explore various types and levels of skills within the spelling, reading and arithmetic domains. For the sake of comparability with other work with these measures, we have used the publisher's scoring procedures, even though some subtest scores therefore represent relatively "pure" skills and others subsume two or more major skill areas within one subtest score.

---

[1] Reliability and validity discussions are in Appendix E.

[2] Subtests are discussed in Appendix C.

The WRAT Spelling subtest, for example, includes mark-copying items, which tap motor development, and also items which call for spelling single syllable words, containing either short or long vowels, or in one case, a diphthong.

The WRAT Reading subtest explores the child's ability to match letters (a visual discrimination task); the ability to name specific letters (a knowledge task); and the ability to read words (which could be either a simple or a complex task: simple if the words are in the child's sight vocabulary, but requiring diverse analytic skills if he must decode them phonically; the items include several types of sounds: long and short vowels, r-controlled vowels, vowel and consonant digraphs, ending consonant blends, and a diphthong).

The WRAT Arithmetic subtest explores the child's ability to count concrete objects, to read digits, to differentiate orally "more" and "less", to perform oral additional and subtraction problems, and to take a written test in simple addition and subtraction.

The PPVT, a vocabulary test, examines the child's ability to choose the appropriate picture for each word enunciated by the examiner.

In the MAT, Primer Level, administered to the kindergarten children in the Spring of 1972, the Listening for Sounds subtest requires that the child match the examiner's vocalization of a sound with the correct picture. Prerequisite skills include a fairly high level of auditory discrimination of beginning consonant and vowel sounds. The second portion of this subtest requires that the child mark either the beginning or ending sound of a pronounced word; this task requires not only fine auditory discrimination, but also specific knowledge of sound-letter correspondence. In the third portion of the subtest, the child must identify the pronounced word among a group of similar words (e.g., one, no, on, nor, or ran, fan, pan, tan). Some of these items could indicate a child with reversal problems (e.g., no, on); others tap his ability to differentiate among very similar words out of context. In general, Listening for Sounds taps skills which are emphasized by phonic and linguistic approaches to reading.

The MAT Reading subtest, on the other hand, requires the child to apply a variety of skills as he identifies upper and lower case letters and matches words with pictures. He might arrive at a correct answer by any or all of a number of means such as sight vocabulary, phonic skills, elimination procedures, and configuration or context clues. A child who is being taught through a relatively eclectic approach, such as a well taught multifaceted language experience approach, is likely to be developing such a diversity of skills.

The MAT Numbers subtest includes both concept manipulation and computational skill. With only one Numbers score, however, it is not possible to distinguish coexisting levels of these component skills. Similarly, the Arithmetic subtest of the MAT at the first grade level also draws upon both concept manipulation and computational skills.

The subtests of the first grade MAT which explore a child's ability to read, extend the pattern of the kindergarten MAT. The Word Analysis subtest is a more difficult version of Listening for Sounds. Here the child must select the pronounced word from a list (e.g., sail, pail, nail, fail, or could, should, would, ward). The separate subtests, Word Knowledge and Reading, which together comprise the Total Reading score, continue and extend the exploration of skills as the kindergarten MAT Reading subtest. In the Word Knowledge subtest, the child must identify the word which matches each of a series of pictures. The Reading subtest is divided into two sections. In the first section, the child must mark the sentence which matches the picture. In the second section, the child must read a short paragraph and answer questions about it; this section adds comprehension to the extensive list of potential skills.

In general, the various reading subtests at both levels of the MAT (kindergarten and first grade) probe a variety of skills, some emphasizing one major skill--phonic decoding--and others encouraging a multiplicity of approaches. Since different reading approaches emphasize different skills during the early reading phase, a child taught by a given approach is likely to have greater strength in one or more types of skills than in others, and is likely to perform differentially on the various subtests. At the same time, because of the diversity of skills tested and scored separately, it is fairly likely that at least one subtest score will reflect a given child's initial reading skills strengths.

CHAPTER IV

COVARIANCE SELECTION FOR THE SPONSOR WITHIN FT/NFT ANALYSES

4.0 INTRODUCTION

Extraneous factors may confound the assessment of an experimental program. When such confoundings are present, the outcomes of the experimental progam may contain the effects of these extraneous factors and their interactions with the treatment as well as the effect of the experimentally defined treatment.

Within a controlled experimental situation, the treatment is rigorously administered to ensure maintenance of the specified treatment across time and subjects, and the differential influence of extraneous factors across treatment groups is controlled through random assignment of subjects to treatment groups.

When these two procedures cannot be implemented, as is often the case in educational research when an experiment is located in a natural setting, a statistical technique, usually the analysis of covariance, may be introduced in an attempt to adjust for the influence of these extraneous factors. Several types of external factors may potentially influence the outcome of the FT experiment. We considered the following categories of extraneous factors for potential inclusion in our set of covariates: geographic and demographic characteristics, school district variables, and characteristics of the schools, classes, teachers, parents and children in our subset. These categories were operationally defined in our Interim Report. Our first step was to examine the distributional properties of all the variables originally proposed. If the distribution was approximately normal, further study of that covariable began immediately--with no change in the variable definition. In all other cases, before the analysis resumed the variable was either transformed (monotonically) or coded into categories having roughly equivalent frequency counts.

Here we present the development of the final covariate set. It should be noted that only four covariables survived the various checks described. The inability to utilize some of the potential variables as covariables may have left some bias in the analyses. However, the covariables utilized, described in Section 4.0, include both socioeconomic and achievement variables.

4.1 INITIAL COVARIATE SET

The final operational definitions of the potential covariates were as follows:

Geographic and Demographic Variables

- Region. Dummy coded variables representing four regions of the country, namely, Northeast, North Central, West, and South.
- City size. The city or town's population as recorded in the 1970 census count and coded into four categories:
    (1) rural area, less than 10,000
    (2) small city, between 10,000 and 50,000
    (3) medium city, between 50,000 and 200,000
    (4) large city, greater than 200,000.
- Density. The number of persons per square mile in the city or town as recorded in the 1970 census count. This variable was dummy coded into two categories:
    0 = less than 4,000 persons per square mile
    1 = more than 4,000 persons per square mile

School District Variables

The following information was obtained from the 1970 Census Fourth Count School District Data Tapes.

- Place description code. Dummy coded variables describing the location of each school district in terms of:
    -- central city of SMSA
    -- central city of urban area
    -- central city of SMSA and urban area
    -- incorporated place
    -- unincorporated place
    -- not in a place
- Elementary enrollment. Number of pupils in public elementary schools in the district.
- Percent Spanish-speaking in district.
- Percent non-English-speaking in district.
- Percent families in district below poverty level.
- District mean income level.
- District mean education level of persons 25 years and up.
- Occupational index for district.

School Variables

- Title I funds per pupil. Obtained from the school questionnaire and dummy coded into two categories:

0 = less than $50 per pupil

1 = greater than $50 per pupil

- Person days per week per pupil of support services. Namely, psychological, medical, instructional, and social services as listed on the school questionnaire.
- School size. The number of students in the school as listed on the school questionnaire.
- Grade span. The number of grades taught in the school as listed on the school questionnaire.
- School integration, from the Directory of Public Elementary and Secondary Schools in Selected Districts, Enrollment and Staff by Racial/Ethnic Group, prepared by the Office for Civil Rights (OCR). The data were coded into three categories:

  (1) more than 90 percent minority (less than 10 percent White)

  (2) between 10 percent and 90 percent minority (integrated)

  (3) less than 10 percent minority (more than 90 percent White)
- Ratio of number of students to number of full-time teachers, excluding part-time equivalents. Computed using the OCR Directory.

Characteristics of the FT/NFT Children in the School

- Entry level. For 1971-72, the mean and standard deviation were computed on all kindergarten Fall WRAT tests available in a school; as an estimate of the 1972-73 entry level of first graders the Spring WRAT tests were also aggregated within a school.
- Mean income, education, and occupational level. The mean of all entering parent interview data (1971-72) for the school, as measures of socioeconomic status.

Class Variables

- Entry level. Mean and standard deviation of the Fall and Spring kindergarten WRAT for all the tested FT/NFT children in the classroom.
- Ethnic composition. Categorized as school integration described above.
- Class size. The number of children in the classroom as reported on the teacher questionnaire.
- Mean income, education, and occupational level. An aggregation of all available parent interview data for both the kindergarten and first grade classes.

- Percent of children eligible for FT.

Teacher Variables

- Experience, education, ethnicity, as reported on the 1971-72 and 1972-73 teacher questionnaires.

Parent Variables

The following information was obtained from the entering Parent Interview (1971-72).

- Index of mother's education, coded as follows:
    - 1 = no schooling
    - 2 = K through 3rd
    - 3 = 4th through 6th
    - 4 = 7th through 8th
    - 5 = high school incomplete
    - 6 = high school complete
    - 7 = college incomplete
    - 8 = college complete
- Index of father's education, coded same as mother's education.
- Income. Total dollars, coded into four categories:
    - (1) less than $4,000
    - (2) between $4,000 and $7,000
    - (3) between $7,000 and $10,000
    - (4) greater than $10,000
- Number of persons in the household. The square root transformation was used because of distributional properties.
- Occupation of head of household, coded as follows:
    - 1 = laborers (farm and nonfarm)
    - 2 = service workers
    - 3 = operatives and kindred workers
    - 4 = craftsmen, foremen and kindred workers
    - 5 = sales, clerical and kindred workers
    - 6 = proprietors, managers and officials, farmers, farm members
    - 7 = professional, semi-professional
- Family structure, coded as:
    - 3 = mother and father present
    - 2 = mother present, but not father
    - 1 = all other possibilities

- Locus of control. As measured in the Fall of 1971, reflecting the extent to which parents consider themselves responsible for events in their lives.
- Parent perception of school receptivity. As measured in the Fall of 1971, reflecting the extent to which parents think they can affect the schools.

Child Variables

- Ethnicity.
- Preschool experience, coded as:
  - 1 = Head Start or any other preschool experience
  - 2 = no preschool experience (stayed home)
- Sex.
- Age, in months.
- Entry level. From the testing in the Fall of the kindergarten year (Fall 1971), there were four test scores that were considered as general covariates for the analysis of the kindergarten data.
  - the Wide Range Achievement Test (WRAT)
  - the Peabody Picture Vocabulary Test (PPVT)
  - the Preschool Inventory (PSI)
  - the Lee - Clark Reading Readiness Test

  From the testing in the Spring of the kindergarten year (Spring 1972), there were two test scores that were considered as general covariates for the analysis of the first grade date.
  - the Wide Range Achievement Test (WRAT)
  - the Peabody Picture Vocabulary Test (PPVT)

4.2 CONSTRAINTS PLACED UPON FINAL COVARIATE SET

These covariates provided the universe from which we made our final covariate selection. Each of the potential covariates was examined in order to ascertain that its usage was not precluded by any of the following constraints:

Lack of confounding with design. Each Sponsor FT/NFT group must be present in each category for the specified categorical covariate. For example, each Sponsor FT/NFT cell must be represented in each defined region of the country. If this is not the case, further statistical manipulation of the variable is precluded.

Presence of data. The operationally defined variable must be present in the data base or readily available from an external source. For example, the percent of children in a class who were eligible for FT was not available.

Correlation between outcome and covariate(s). Covariates isolate those variations in the outcome measure which are independent of treatment effects, and thus expose "pure noise" in the data. The larger the correlation between an outcome and a covariate, the more the covariate will explain the non-treatment-related outcome variations. If the correlation (of outcome and covariate) was less than .20, the covariable was dropped. Teacher's experience and parent perception of receptivity were among the many covariates dropped for this reason.

Correlation between covariates. A covariate was not entered in the model if it correlated as highly with another covariate in the model as it did with the outcome measures. For example, the occupation of the head of household was dropped as a covariate because of its correlation with family income. Often this constraint was examined by the use of partial correlations.

Sample Size. When analyzing any real life data, one expects to find some missing observations; the extent of missing data being usually less than 5 percent, at most 10 percent. In the Follow Through world, however, missingness of data seems to be a rule rather than exception. Thus, when we tried to work with a "pure" sample, i.e., the set of children who had scores on all outcome measures of interest and corresponding policy relevant covariates, we ran the risk of working with an unacceptably small analytic sample. We have, therefore, adopted the strategy of eliminating from further analysis any covariate with substantial amount of missing scores. For example, Parent General Locus and family structure have been eliminated for this reason.

Mother's education and family income were deemed too important to be dropped because of small sample sizes. There also seems to be information in the nonresponse of a parent to questions dealing with these two important SES measures. Missing data codes were introduced for mother's education and family income; (0) indicates a response was given and a (1) signifies no response. The within group (Sponsor by FT/NFT) mean was used on the original covariable (e.g., mother's education and family income) for instances of nonresponse.

Conceptual grounds. The following variables were dropped from the covariate set because they lacked conceptual integrity.

- There was no simple but meaningful way of creating an index of the occupational structure of the district or the school.

- Class variables for entry level, mothers' education and income in favor of those indices aggregated for all sample children in the school. The class and school level variables were largely accounting for the same variance in outcome measure and it was felt that school aggregated measures were:
    - more stable and hence a better estimate of the learning environment
    - more appropriate for use in the school level analysis
    - more closely related to past research (in particular, Coleman's study)
- Father's education was missing for a large percentage of the sample. This missingness could be a function of either the respondent not knowing the father's educational level or the father not being home. When the father's education was available, it was highly correlated (.64 in NFT world) with the mother's education. Previous research has shown that for lower income children mother's education is more related to school performance than is father's education. Hence father's education was removed from the covariate list in favor of mother's education.

Nonsingularity of correlation matrix. If the correlation matrix is singular, further statistical manipulation of the data within the framework of the analysis of covariance cannot proceed. A singular matrix may be the result of zero cell means or colinearity. The school mean WRAT and missing data code for family income are two of the covariates that were associated with a non-invertable (singular) matrix. It was felt that school mean WRAT was functionally colinear with child level WRAT scores (even though the aggregation was based on a slightly larger number than our final ANCOVA analytic sample). The missing data code resulted in a zero mean for some Sponsor treatment groups who had complete data available for family income.

Homogeneity of Regression. Finally, the covariates were checked against all outcome measures for homogeneity of regression. Homogeneity is discussed in Appendix B. We can estimate Sponsor effects in the ANCOVA design when when the Sponsor X Treatment X Covariate interaction terms are nonsignificant. Hence this is the term of interest in our tests for homogeneity.

The next section will make evident that only a few covariables reached this final step in the process of selecting the covariate set. An even smaller number made it through the homogeneity tests. The final covariables are presented in the last section of this chapter.

4.2.1 Summary of Excluded Covariates

We now present the complete list of potential covariates once more. Beside each covariate is listed the reason for its exclusion.

POTENTIAL COVARIATES	REASONS FOR EXCLUSION
Geographic and Demographic Variables	
Region	Confounded with design
City size	Confounded with design
Density	Too highly correlated with other covariate (school integration)
School District Variables	
Place Description Code	Missing Data
Elementary Enrollment	Lack of correlation with dependent measures
Percent Spanish-speaking in district	Lack of correlation with dependent measures
Percent of non-English speaking in district	Lack of correlation with dependent measures
Percent families in district below poverty level	Lack of correlation with dependent measures
District mean income level	Lack of correlation with dependent measures
District mean education level of persons 25 years and up	Lack of correlation with dependent measures
Occupational index for district	Conceptual grounds
School Variables	
Title I funds per pupil	Small sample size
Person days per week per pupil of support services	Lack of correlation with dependent measures

POTENTIAL COVARIATES	REASONS FOR EXCLUSION
School size	Lack of correlation with dependent measures
Grade span	Lack of correlation with dependent measures
School integration	Singular matrix; Heterogeneity of regression[1]
Ratio of number of students to number of full-time teachers	Lack of correlation with dependent measures
Characteristics of the FT/NFT Children in School	
Fall WRAT:	
mean	Singular matrix
standard deviation	Too highly correlated with other covariate (school Fall WRAT mean)
Spring WRAT:	
mean	Singular matrix
standard deviation	Too highly correlated with other covariate (school Spring WRAT mean)
Mean family income	Singular matrix
Mean educational level of mothers	Singular matrix
Occupational index	Conceptual grounds
Class Variables	
Fall WRAT: mean and standard deviation	Conceptual grounds and too highly correlated with other covariate (School level data)
Spring WRAT: mean and standard deviation	Conceptual grounds and too highly correlated with other covariate (School level data)
Ethnic composition	Small sample size
Class size	Missing data

---

[1] After culling those Sponsors who caused the matrix to be singular, we found heterogeneity of regression.

POTENTIAL COVARIATES	REASONS FOR EXCLUSION
Mean family income	Conceptual grounds and too highly correlated with other covariate (School family income)
Mean educational level of mothers	Too highly correlated with other covariate (School Mother's education)
Occupational index	Conceptual grounds
Percent of children eligible for FT	Missing data
Teacher Variables	
Experience	Lack of correlation with dependent measures
Education	Lack of correlation with dependent measures
Ethnicity	Too highly correlated with other covariate (school integration)
Parent Variables	
Education level of mother	Heterogeneity of regression (with first grade outcomes)
Missing data code for education level of mother	Singular matrix
Education level of father	Conceptual grounds; small sample size; too highly correlated with other covariate (Education level of mother)
Family income	Heterogeneity of regression (with first grade outcomes)
Number of persons in household	Lack of correlation with dependent measures
Occupation level of head of household	Too highly correlated with other covariate (income)
Family structure	Small sample size
(General) Locus of Control	Small sample size and lack of correlation with dependent measures
Parent perception of school receptivity	Lack of correlation with dependent measures

POTENTIAL COVARIATES	REASONS FOR EXCLUSION
Child Level Variables	
Fall WRAT	Heterogeneity of regression (with some first grade outcomes for some Sponsors)
Fall PPVT	Too highly correlated with other covariate (Fall WRAT)
Fall PSI	Too highly correlated with other covariate (Fall WRAT)
Fall Lee-Clark	Too highly correlated with other covariate (Fall WRAT)
Ethnicity	Confounded with design
Preschool experiences	Lack of correlation with dependent measures
Sex	Lack of correlation with dependent measures
Age	Lack of correlation with dependent measures

4.3 FINAL COVARIATE SET

After careful examination of all the potential covariates, only three[2] survived for the overall analysis of Sponsor FT/NFT differences: the WRAT as administered in the Fall of the kindergarten year, family income, and mother's education. For the reanalysis of the kindergarten outcomes (as done last year) the Fall WRAT, family income and mother's education were used as covariates; for the analysis of first grade outcomes the Fall WRAT was the only covariate. Because of the heterogeneity of the FT and NFT relationship between end of first grade outcome measures and the kindergarten Fall WRAT, adjusted estimates of the FT/NFT difference could not be computed. The details of this heterogeneity are presented in the two-year effects chapter.

---

[2] The Fall PPVT also survived; however, it was used, in place of the Fall WRAT, only as a covariate for analyzing the Spring kindergarten PPVT.

CHAPTER V

KINDERGARTEN STUDY: A REPLICATION

5.0 INTRODUCTION

There are two major purposes of this chapter. The first is to describe the two year traced sample utilized in the kindergarten analyses, and to compare this sample with the 1973 analytic sample on which the kindergarten analyses were performed last year and reported in Volume IA of this series. The process of tracing changed the sample to some extent and this may influence the analytic findings.

The second purpose of this chapter is to report the results of the one year traced sample kindergarten study as a replication of the same study reported in the earlier volume.

Since the primary purpose of this chapter is to report the effects of the two years of continuous FT experience with the same Sponsor, it was necessary to identify a group of FT and NFT children who were present at the three testing points included in the two year, kindergarten-through-first grade span. These points are the fall of 1971 when Cohort III children entered kindergarten; the Spring of 1972 when these children completed kindergarten; and the spring of 1973 when the children completed first grade. The children who were present at all three testing points and who satisfied the other inclusion criteria (see Chapter II) are called the traced, 1974, analytic sample. The children who were present during the first two testing points, and for whom appropriate sets of data are available, constitute the 1973 analytic sample on whom last year's kindergarten study was based. A detailed comparison of these two groups of children is contained in Chapter II of this Volume.

The traced sample is a subset of the 1973 sample and thus ordinarily contains fewer children than the latter. There are several reasons why these two are not identical.

1. Some children left the Follow Through experiment after kindergarten. This may have occurred because their families left the area, or otherwise decided to withdraw the children from the program. It may also have occurred because the children were assigned to other schools or classes by the school administration. There is anecdotal evidence that some school administrators removed higher performing children from FT classes and schools to make room for what the administrators considered to be more

deserving children. Other administrators removed lower achieving children because the children were assumed to be hindering the operation of the program. Children who were sent into the FT classes as replacements at the beginning of first grade are not included in these analyses since they do not have fall, 1971, testing data (pretest) or spring, 1972, testing data (post-kindergarten).

2. Some children were absent for testing at the end of first grade or their tests were invalidated because of a testing anomoly .

3. Some children were left out of the testing because the end of first grade is, in the FT national evaluation plan, considered to be an interim testing point at which only a subset of the kindergarten graduates were scheduled for testing.

The net result of these losses is that the traced sample is considerably smaller than the 1973 sample, although for some Sponsors there was no appreciable loss. Since the net loss does not appear to be random (see Chapter II of this Volume), the magnitude of the loss for each Sponsor must be considered as a reflection of the unique factors at each site. Unfortunately, these factors cannot be identified or quantified, and must remain as a source of confounding of Sponsor effects.

There is one other factor which contributes to the difference between the traced and the 1973 samples. In order to reduce analytic complexity, all children who were not Black or White were excluded from the 1973 sample. The criterion was not applied to the traced sample. A substantial number of Spanish-surnamed, Oriental, and American Indian children is included in the traced sample, which reduces the numerical difference between the two but increases the substantive differences. This addition is also not random. For example, High Scope, Florida, and SEDL gained large numbers of Spanish-surnamed children, and the remaining Sponsors did not.

A summary of the net change for each Sponsor is presented in Table V-1. It is clear that Kansas, High/Scope and SEDL have essentially the same number of children in both analyses. High/Scope and SEDL have these numbers because they lost a great many children through tracing, but gained (primarily Spanish-surnamed) children through the change in the inclusion criteria. Kansas lost many children as a result of the tracing, but gained many children because the criterion of a full set of data on mother's education which was applied to the 1973 sample was not applied to

TABLE V-1

DISTRIBUTION OF 1973 AND 1974 FINAL ANCOVA SAMPLES FOR CHILD LEVEL STUDIES BY SPONSOR BY FT/NFT

Sponsor	1973 Sample		1974 Sample	
	FT	NFT	FT	NFT
Far West	302	303	219	178
Arizona	317	265	260	166
Bank Street	212	131	119	67
Oregon	153	119	95	82
Kansas	306	110	320	113
High/Scope	173	144	168	145
Florida	243	101	173	74
EDC	250	147	150	96
Pittsburgh	258	160	153	60
SEDL	75	61	67	50
Total	2289	1541	1724	1031

the traced sample (a missing data code was used for the traced sample). Thus the Kansas sample includes many children whose mothers did not complete their interviews.

All remaining Sponsors showed a net loss, but Bank Street and Pittsburgh lost as much as two thirds of the children in some categories. These represent losses of entire sites for both of these Sponsors, which leaves a sample that cannot be considered at all close to the population with which these Sponsors are working.

Before turning to the description of the samples and the analytic findings, it is appropriate to summarize the analytic methods used and to explain the method of presentation of findings. These methods of analysis and presentation are identical to those in the analyses reported in Volume IA, the first volume in this series. The reader familiar with that report should forego the following two sections.

Analytic Techniques

Two statistical techniques are utilized in these analyses:

1. Analysis of Variance which yields the unadjusted raw difference between FT and NFT means on an outcome measure, and

2. Analysis of Covariance which yields an adjusted effect, the difference between the FT and NFT outcome means adjusted for differences between the FT and NFT on MOTHERS' EDUCATION AND FALL, 1971 ENTRY TOTAL WRAT SCORE. These two measures are the covariables used throughout these analyses.

Note that an "effect" as used here means a difference between FT and NFT means. An adjusted "effect" means a FT/NFT difference in outcome means adjusted for differences between these groups on the two covariables. The FT and NFT are then examined for their outcome differences "as if" they started out with equal entry test scores and SES, as measured by levels of mother's education.

For a full description of the statistical procedure the reader is referred to Appendix B of this volume.

Method of Presentation of Results

For each Sponsor a discussion of the similarities and differences between FT and NFT groups on demographic characteristics and entry test scores is presented, followed by a comparison of 1973 and 1974 analytic samples and a discussion of one-year kindergarten effects for the present analytic sample. Figure V-1 illustrates the content of all effects tables presented in this volume. The table displays the results of ten parallel analyses which yield measures of the difference between FT and NFT on ten criterion measures. The statistics tabulated at the bottom of the figure document the derivation of the graphical display. They include:

- Adjusted and unadjusted values of the main effects, expressed in the units of the criterion variables. The unadjusted WRAT spelling effect of -1.649 (in the leftmost column) implies, for example, that the FT group scored lower than NFT. With mother's education and entry WRAT inequalities taken into account by covariance adjustment, NFT's advantage decreases to -.855 points. The adjusted effects are computed by ANCOVA and the unadjusted by ANOVA. We refer to these effects by the algebraic symbol B because we have selected as a computational vehicle for our analyses, a particularly flexible multiple regression formulation which yields the "effects" of ANOVA and ANCOVA as raw score regression weights ("B weights").

- The standard errors of the adjusted and unadjusted main effects, which we compute.

- F ratios as the ratios of the effects to their standard errors squared. As Appendix B on methodological issues makes clear, one can regard the squares of these F ratios as scaled signal-to-noise ratios reflecting the extent to which the observed patterns of effects emerge clearly from the undifferentiated criterion variance that would otherwise seem unrelated to the treatments and covariables.

- The standard deviation of the criterion, reflecting the variability of each outcome variable from school to school within the population under analysis. We introduce this statistic only so as to be able to compare the magnitudes of effects from one outcome variable to another and use the criterion of a difference greater than .25 standard deviations as indicative of an educationally significant difference.

- The effects expressed in criterion standard deviations, permitting the reader to say, for example, that FT's overall adjusted "effect" on the reading subtest of the Wide Range Achievement Test (WRAT) scores amounts to 0.411 standard deviations. The arrows of the graphical displays correspond to these standardized effect measures, both adjusted (solid arrows) and unadjusted (dashed arrows).

- The number of children used in the computation. This number has a strong influence, of course, on the standard errors of the effects and therefore on the F ratios. With large N's, it is possible to have large F's that correspond to totally uninteresting effects; with small N's, on the other hand, really important and revealing patterns can wash out in the noise.

5.1 FAR WEST, RESPONSIVE EDUCATIONAL PROGRAM

5.1.1 Sample Description

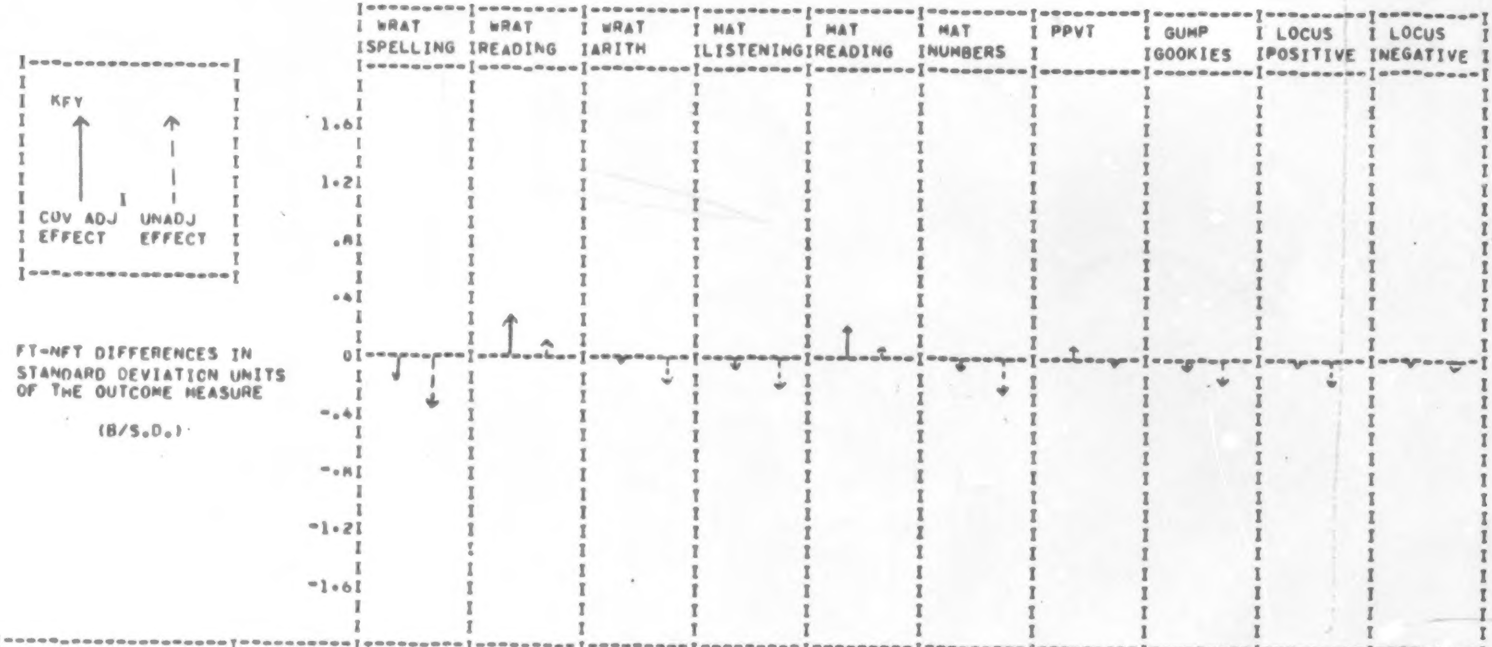
The FT and NFT samples for Far West Laboratory have similar backgrounds although the FT group appears to be somewhat more economically disadvantaged. The FT families report a lower level of income, but the FT/NFT mothers have similar educational backgrounds, and only a slightly larger proportion of the FT children come from single parent homes. The entry level scores for these two samples indicate this same pattern of fairly good matching. They have similar entry level PPVT scores, but the FT group received lower scores on the fall WRAT. The FT sample also had a larger proportion of children who attended preschool.

This year's FT and NFT samples are located in large and medium-sized cities. About one-quarter of them are located in large Northeast and North Central cities and the remainder in medium-sized North Central and Western cities. Last year's sample was distributed in a similar fashion.

Last year's exploration of FT/NFT mismatch indicated that this Sponsor's FT and NFT samples were among the most evenly matched. Comparing this year's and last year's samples, we find that this year's FT sample is somewhat poorer and lower on entering achievement scores than last year's while the NFT sample is similar to last year's NFT sample. Hence, Far West's current FT/NFT mismatch, albeit slight, is greater than that of last year.

5.1.2 Kindergarten 1-Year Effects

Fig. V-1 indicates that there is only one area of meaningful FT/NFT contrasts. Both the MAT and WRAT Reading subtests show FT scores higher than NFT scores. This is very similar to the kindergarten effects found in the original analyses performed on the 1973 Analytic Sample except that the current analysis based on the traced subset of the full group shows somewhat stronger reading effects. The current adjustment for entry differences between FT and NFT is almost a full point on the WRAT Reading subtest and reflects the relatively lower entry scores of the FT group. Since the FT traced group is somewhat lower on entry scores than the 1973 Analytic Sample, and since the reading effects are a little stronger in the traced group, it is reasonable to suspect that the Far West program



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.855	1.619	-.008	-.453	1.125	-.560	.926	-.015	-.094	-.008
	UNADJ	-1.649	.625	-.698	-1.568	.339	-1.703	-.361	-.028	-.351	-.094
S.E. = STANDARD ERROR OF B	ADJ	.380	.428	.283	.590	.414	.479	.609	.015	.193	.164
	UNADJ	.456	.567	.364	.697	.505	.596	.900	.016	.204	.165
F= SIGNIFICANCE STATISTIC	ADJ	** 5.068	** 15.638	.009	.589	** 7.378	1.368	2.311	.991	.239	.002
	UNADJ	** 13.084	1.214	* 3.669	** 5.065	.450	** 8.173	.161	* 3.151	* 2.949	.328
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.182	.278	-.002	-.062	.218	-.088	.098	-.096	-.046	-.005
	UNADJ	-.352	.197	-.183	-.216	.065	-.269	-.038	-.178	-.172	-.058
N	FT	219	219	219	219	219	219	219	219	219	219
	NFT	178	178	178	178	176	178	178	178	178	178

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

FIGURE V-1 KINDERGARTEN EFFECTS PROFILE FOR FAR WEST ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

may be more effective for the lower achieving children than for the higher achieving children. Although definitive data are not available to examine this possibility, it is appropriate to refer to the kindergarten effects reported by site (Figures C -1 through C-5, Appendix C) where a great deal of variability over these sites is apparent. The reading effects are found in two of the five sites, and these sites tend to be in the lower income, lower achieving range. Indeed, the only site showing consistently strong FT effects (NC2) is relatively low on both family income and entry scores. However, these findings do no more than make this explanation of the relationship between the effects and the mismatch worth exploring in greater detail. The alternative explanation, that the nature of the mismatch varies by site, and that the site characteristics account for most of the important findings for this Sponsor, will be explored when the site data are presented with the two year findings in the next chapter.

In general, the Far West children are holding their own with respect to NFT. However, the program does not appear to produce a distinctive pattern of effects in terms of outcome domain.

5.2 ARIZONA, TUCSON EARLY EDUCATIONAL MODEL

5.2.1 Sample Description

This year's traced FT and NFT samples for the University of Arizona, like last year's, are severely mismatched. The mothers of the FT children are less well educated than the mothers of the NFT children. The NFT sample's mean income is the highest reported for any group in this year's analysis, whereas the FT sample's mean income is substantially lower. The FT sample children also come from larger families. These socioeconomic indicators of poorly matched samples are paralleled by the differences in performance level of the FT and NFT children at entry. The University of Arizona is the only Sponsor whose FT sample scored below the NFT sample on both the entry level PPVT and WRAT. Our only indicator on which this FT sample exceeded the NFT sample was preschool experience, reflecting selective Head Start participation by disadvantaged children.

Nearly half of this year's FT/NFT sample resides in large cities located in the Northeast and North Central regions. However, 21 percent of the FT children live in Northeastern large cities while fewer than ten percent of the NFT children do. This larger proportion of Northeastern large

00095

city children may be reflected in the larger proportion of Blacks in the FT sample. The non-big-city children in Arizona's samples are located primarily in the Northeast and North Central regions with only 11 percent (FT) and 13 percent (NFT) in the South.

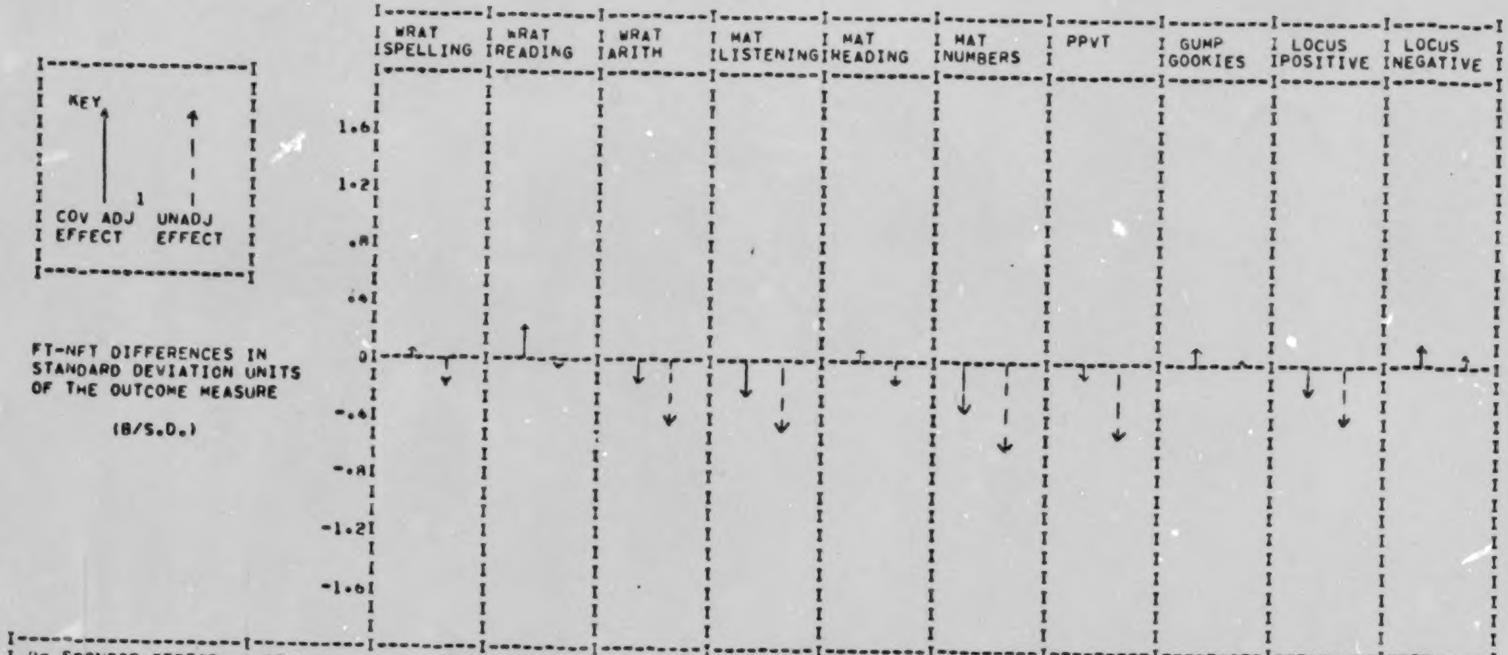
Last year's sample was distributed across city size in a pattern quite similar to this year's, but fewer of last year's sample children live in the Northeast and more live in the South. The FT/NFT differential geographic distribution is similar across the two samples.

5.2.2 Kindergarten 1-Year Effects

The FT/NFT differences for Arizona, shown in Figure V-2 , indicate an NFT advantage on the MAT Listening for Sounds and Numbers subtests, duplicating last year's findings.

Despite substantial adjustments to account for the large initial differences between the FT and the NFT, the children participating in the Arizona model are still meaningfully behind the NFT children on the WRAT Arithmetic, MAT Numbers and Listening for Sounds, and PPVT measures. Only in the reading domain does the FT treatment contribute as much as the NFT to the children's achievement. In fact, the WRAT Reading subtest shows an FT effect which almost reaches significance. This is also true for measures of academic motivation (Gumpgookies) where the FT group tends to show a greater effect. This pattern of effects appears to be consistent across the several Arizona sites and the 1973 and 1974 Analytic Samples.
The initial differences between the FT and the NFT groups appear to be so great that the adjustment techniques used in these analyses cannot reveal the full extent of the Arizona contribution. It is only in the reading and motivation measures, where the FT and NFT children are approximately equal on entry into kindergarten that the Arizona program shows a reasonable contribution.

The FT/NFT kindergarten differences for Arizona's sites, shown in Appendix C, Figures C-6 through C-10, indicate consistent and substantial NFT advantages on many measures. FT is behind NFT at the end of kindergarten on the PPVT, on the WRAT Arithmetic subtest, and on the MAT Numbers and Listening for Sounds subtests. On the Reading subtests and on the Gumpgookies measure of achievement motivation FT is at least even with the NFT group.



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP COOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	.250	1.351	-.677	-1.989	.489	-2.196	-1.072	.022	-.437	.252
	UNADJ	-.977	-.405	-1.754	-3.766	-.786	-3.983	-5.138	.002	-.865	.116
S.E. = STANDARD ERROR OF B	ADJ	.377	.424	.281	.585	.411	.475	.605	.015	.191	.162
	UNADJ	.449	.559	.359	.686	.497	.587	.886	.016	.201	.162
F= SIGNIFICANCE STATISTIC	ADJ	.439	** 10.147	** 5.811	** 11.552	1.414	** 21.374	* 3.135	2.036	** 5.224	2.398
	UNADJ	** 4.741	.525	** 23.900	** 30.161	2.504	** 46.110	** 33.635	.010	** 18.495	.517
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.053	.232	-.178	-.274	.094	-.346	-.114	.137	-.214	.154
	UNADJ	-.209	-.069	-.460	-.520	-.152	-.628	-.546	.009	-.423	.071
N	FT	260	260	260	260	260	260	260	260	260	260
	NFT	166	166	166	166	166	166	166	166	166	166

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT THE FALL WRAT WAS REPLACED WITH THE FALL PPVT.

** P LESS THAN .05 * P LESS THAN .10

FIGURE V-2 KINDERGARTEN EFFECTS PROFILE FOR ARIZONA ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

Arizona thus appears to be showing a fairly consistent pattern of effects in kindergarten, producing small or marginal advantages in reading and in achievement motivation, but losing ground in other academic areas.

5.3 BANK STREET COLLEGE APPROACH

5.3.1 Sample Description

The Bank Street traced sample shows an extremely large loss of children from the full 1973 analytic kindergarten sample. Whereas the across-sponsor FT and NFT traced sample is 75% and 67%, respectively, of the 1973 analytic sample, the Bank Street sample is only 56% and 51%. Even more important is the disproportionate loss of children across all locations in which Bank Street operates. These losses are due to direct attrition and/or missing data. Although all of Bank Street's sites are in the Northeast, the loss is not distributed equally across sites of various sizes. Of the 190 children in small cities who were part of the 1973 kindergarten sample, only 45 could be traced for the present sample. Of the 93 children in the large city sample in the 1973 analytic sample, only 45 could be found for the traced sample. The number of children in the medium-sized cities remained constant over the two samples (86 and 76 respectively).

Although the loss of children is only slightly greater in the NFT group than in the FT group, there appears to be a unique loss from the high income, high achieving end of the NFT distribution. Mothers' education changed so drastically from the full sample to the traced subset that the NFT mothers' mean educational level is lower than the FT mothers' mean in the traced sample whereas it was above in the 1973 sample. This change is accounted for primarily by the changes which took place in one of the small cities. A similar reversal of status took place in the distribution of PPVT scores. All of the other SES and achievement measures showed a smaller discrepancy between FT and NFT in the traced sample because the mean NFT scores dropped almost to the level of the FT means. These changes are puzzling since they represent events occurring in the Bank Street sites which reflect none of the expected decisions which have been observed in other sites for other Sponsors. There is no reason to expect such a high rate of loss in the sample or such a peculiar loss among the NFT children. At the same time, there is anecdotal evidence that high achieving FT children were leaving the sample in some sites which suggests that one set of factors was influencing the participation of the FT children and another set of factors was influencing the participation of the NFT children. Under these conditions interpretation of the findings is most difficult, if not impossible.

5.3.2 Kindergarten 1-Year Effects

Figure V-3 presents Bank Street's FT/NFT kindergarten differences. In general, the findings are similar. The FT children show approximately the same adjusted levels as the NFT children on the MAT Numbers and WRAT Arithmetic subtests. The FT children show much lower adjusted levels on all other academic measures. These results cannot be easily interpreted at the overall Sponsor level. Site level analyses presented in the two year study consider the implications of the Bank Street findings in greater detail (See Chapter VI).

5.4 UNIVERSITY OF OREGON

5.4.1 Sample Description

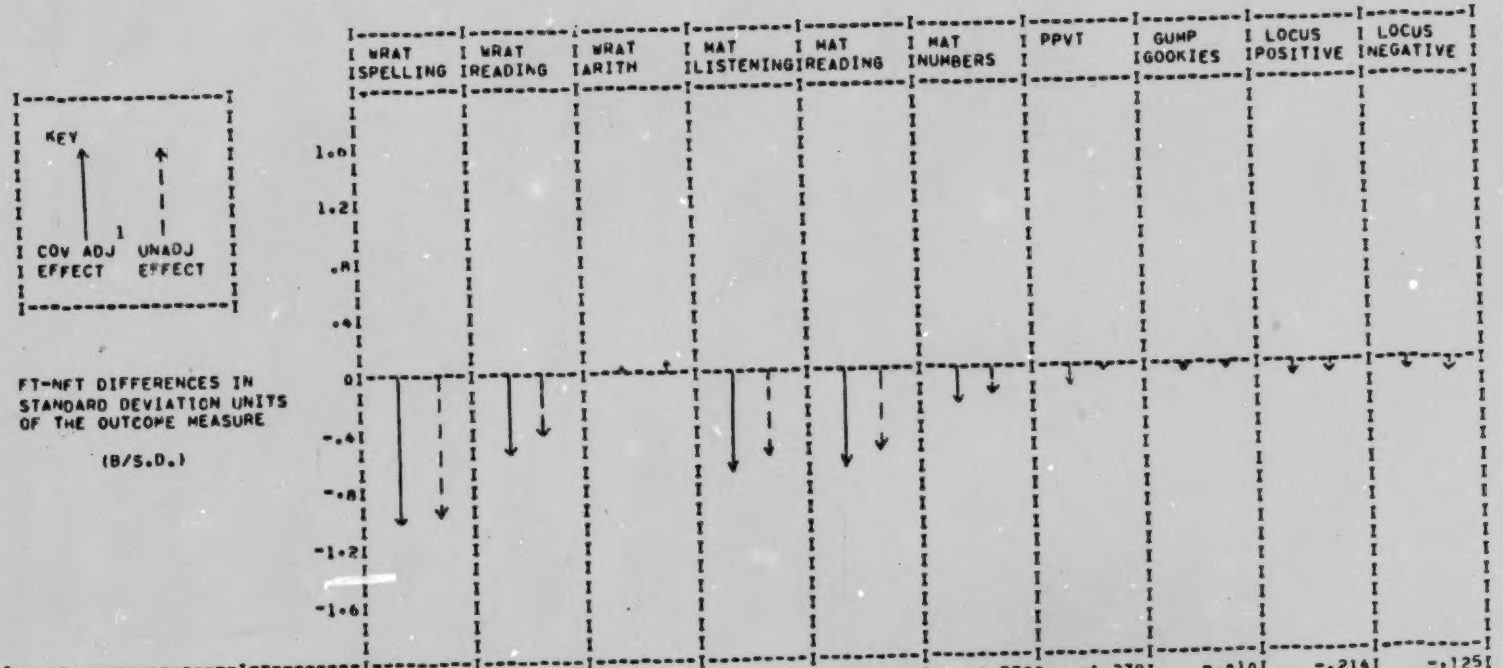
The University of Oregon's FT and NFT samples are severely mismatched on socioeconomic measures. The FT sample has a lower level of income, fewer two-parent homes and less educated mothers than the NFT sample. The two samples are quite similar, however, at entry level on both the WRAT and the PPVT. These findings may in part be related to the fact that all of the FT sample attended some form of preschool while only a very small percentage of the NFT sample did. We hypothesized last year that early treatment effects might explain these findings.

Oregon's sample is located in medium-sized cities in the Northeast and North Central regions. However, while less than ten percent of the FT sample resides in the Northeast, over 20 percent of the NFT sample does. The majority of both samples (92% = FT, 78% = NFT) resides in the North Central region. Last year almost 40 percent of the FT sample resided in the Northeast with a smaller proportion in the North Central region. Last year's NFT sample was distributed quite similarly to this year's.

The mismatch for last year's sample is very similar to that found this year. In both instances socioeconomic FT disadvantage accompanies fairly equal entry level scores.

5.4.2 Kindergarten 1-Year Effects

The end of kindergarten FT/NFT differences for Oregon, shown in



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-4.876	-3.242	.075	-4.790	-3.441	-1.572	-1.378	-.010	-.214	-.125
	UNADJ	-4.637	-2.500	.356	-4.240	-2.873	-1.226	-.286	-.004	-.135	-.122
S.E. = STANDARD ERROR OF B	ADJ	.576	.649	.430	.895	.628	.726	.924	.023	.293	.248
	UNADJ	.690	.859	.552	1.054	.764	.902	1.362	.024	.309	.249
F= SIGNIFICANCE STATISTIC	ADJ	** 71.765	** 24.987	.030	** 28.661	** 29.999	** 4.682	2.222	.176	.536	.254
	UNADJ	** 45.182	** 8.471	. 18	** 16.179	** 14.148	1.848	.044	.031	.190	.240
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-1.041	-.556	.020	-.661	-.666	-.248	-.147	-.062	-.105	-.076
	UNADJ	-.990	-.429	.094	-.585	-.556	-.193	-.030	-.027	-.066	-.075
N	FT	119	119	119	119	119	119	119	119	119	119
	NFT	67	67	67	67	67	67	67	67	67	67

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT.

** P LESS THAN .05 * P LESS THAN .10

FIGURE VI-3 KINDERGARTEN EFFECTS PROFILE FOR BANK STREET ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

Figure V-4 , indicate FT-favoring effects for all outcomes except the PPVT, the Gumpgookies,and Locus Negative, a pattern that mirrors last year's analysis. Despite the initial differences on SES indicators, the Oregon FT children showed significantly higher levels of performance than the NFT children. To be sure, these NFT children had surprisingly low entry scores for their relatively high SES, but the nature of the FT/NFT contrasts for Oregon will be examined further with the individual site data presented in the next chapter. One year kindergarten effects are, for this Sponsor, strongly in favor of the FT group in the traced sample as they were in the 1973 analytic sample.

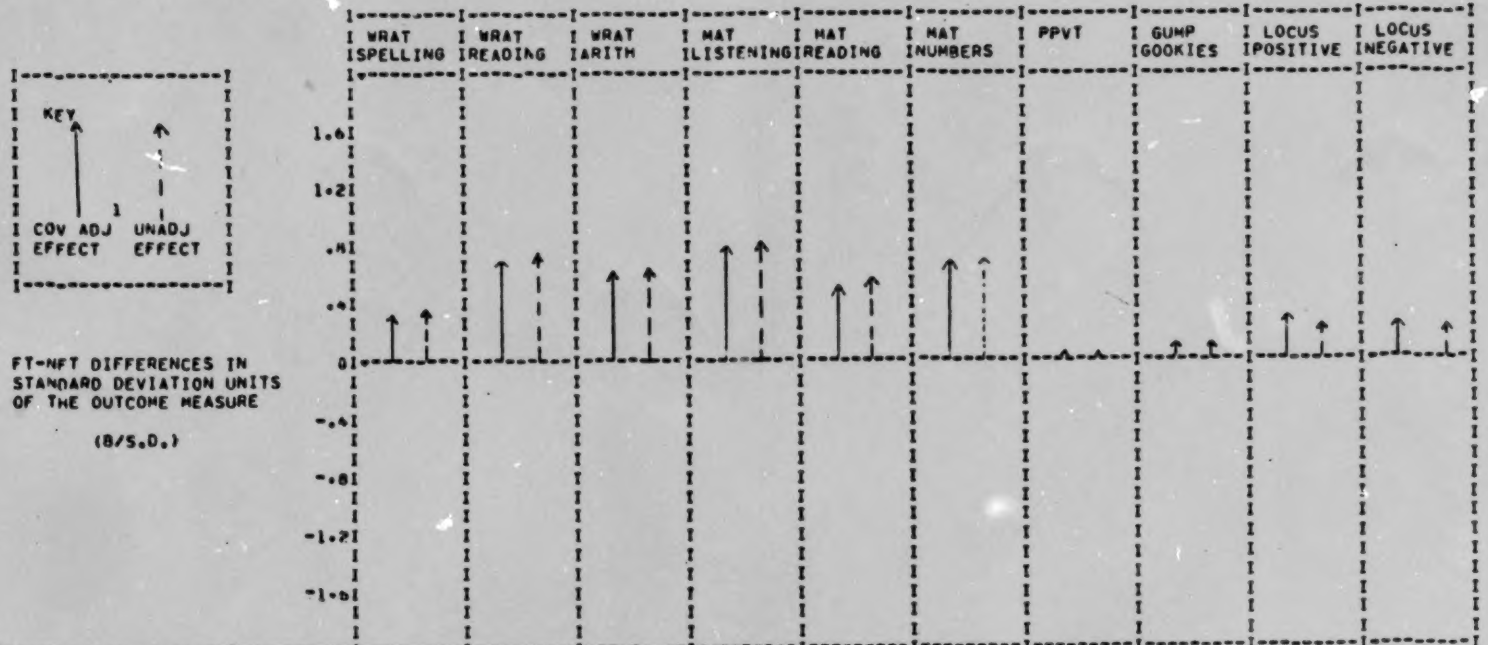
5.5 UNIVERSITY OF KANSAS, BEHAVIOR ANALYSIS APPROACH

5.5.1 Sample Description

Socioeconomically, the FT and NFT samples for the University of Kansas appear to be quite well matched. The FT sample achieved higher scores on the entry level WRAT than the NFT sample, but not on the entering PPVT test. The FT children are over a month younger than the NFT children. Approximately the same proportions of FT and NFT children attended some form of preschool.

A comparison of the 1973 analytic sample with the traced sample of Kansas children shows the two groups to be very similar. The traced FT group is about as far ahead of the NFT group in the 1973 sample. The result of the tracing reduced the PPVT scores in both the FT and the NFT groups by equal amounts. No other meaningful differences resulted from the tracing. The Kansas group of FT children starts out ahead of the NFT children in overall mean WRAT, and this is a factor which will seriously influence the interpretations of findings associated with this Sponsor. The issue will be addressed in Chapter VI where the findings by individual site cast light on the Kansas FT/NFT mismatch.

Geographically, the FT and NFT sites are not distributed similarly. Perhaps the most glaring geographic mismatch is that 34 percent of the FT sample is found in the Northeast large cities, where no NFT pupils reside. Almost half of the Kansas FT sample resides in the Northeast, while only 11 percent of the NFT sample does. Two-thirds of the NFT sample but only one-third of the FT sample resides in the North Central region. Last year's sample distribution was similar to this year's except that there was a



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	1.555	4.022	2.283	5.769	2.612	4.277	.166	.014	.597	.396
	UNADJ	1.634	4.236	2.405	5.939	2.888	4.302	.176	.017	.492	.364
S.E. = STANDARD ERROR OF B	ADJ	.568	.640	.424	.883	.620	.716	.912	.023	.289	.245
	UNADJ	.681	.849	.544	1.040	.754	.890	1.344	.024	.305	.246
F= SIGNIFICANCE STATISTIC	ADJ	** 7.503	** 39.531	** 29.030	** 42.725	** 17.768	** 35.640	.033	.386	** 4.286	2.606
	UNADJ	** 5.761	** 24.976	** 19.530	** 32.593	** 14.670	** 23.365	.017	.496	2.603	2.198
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.332	.690	.599	.796	.505	.675	.018	.090	.292	.242
	UNADJ	.349	.727	.631	.820	.558	.679	.019	.105	.241	.223
N	FT	95	95	95	95	95	95	95	95	95	95
	NFT	82	82	82	82	82	82	82	82	82	82

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

FIGURE V-4 KINDERGARTEN EFFECTS PROFILE FOR OREGON ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

small proportion (6%) of the NFT sample in the Northeast large cities. The proportion of the FT sample in Northeast large cities was slightly smaller last year (28%) than this year (34%).

5.5.2 Kindergarten 1-Year Effects

The Kansas FT/NFT kindergarten differences shown in Figure V-5 indicate FT-favoring effects on all academic outcomes except the PPVT, just as last year. This pattern of substantial achievement advantages accrued during kindergarten on all outcomes except the PPVT appears in three of Kansas' four sites (see Appendix C, Figures C-16 to C-19). The fourth site, NC1, shows substantial FT-favoring unadjusted effects, which were attenuated when the initial advantage of the FT children in this site was taken into consideration. Kansas does appear to produce a consistent pattern of favorable effects.

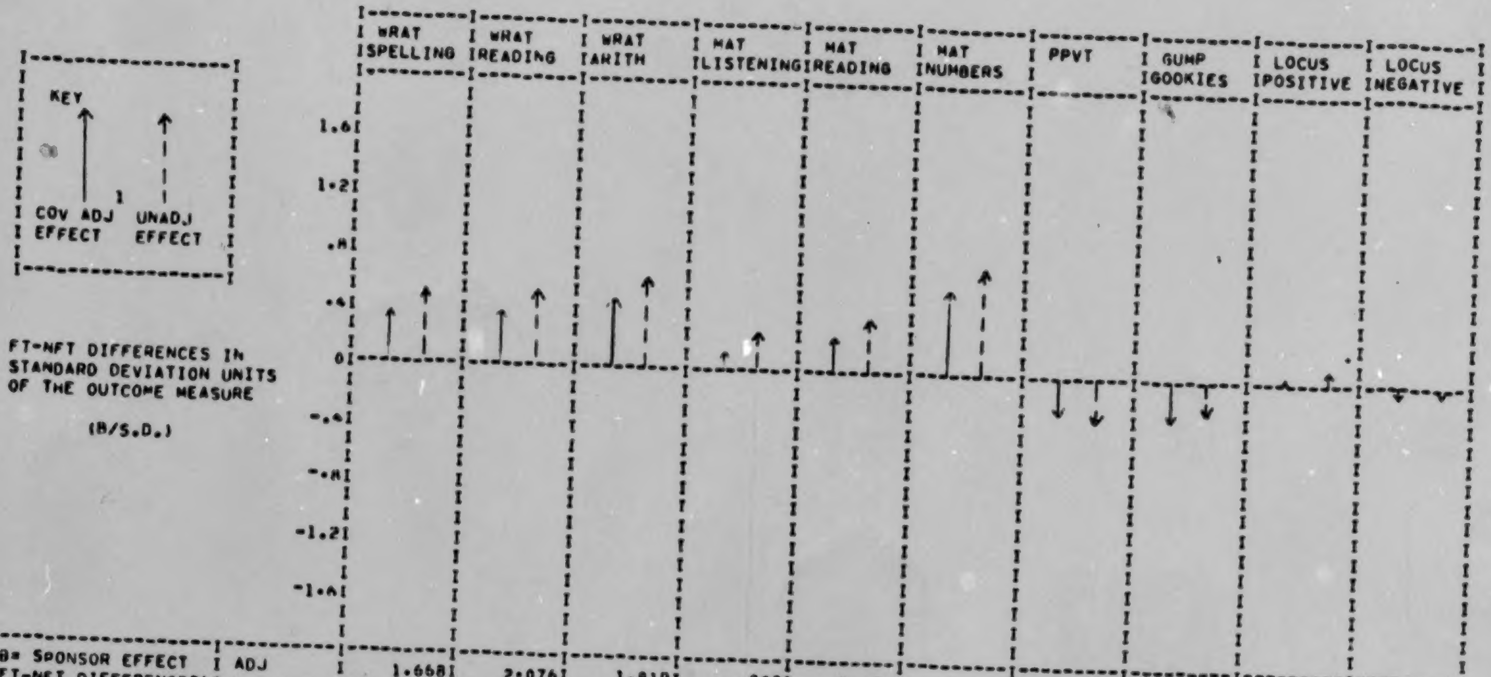
5.6 THE HIGH/SCOPE FOUNDATION, COGNITIVELY ORIENTED CURRICULUM

5.6.1 Sample Description

High/Scope's FT and NFT samples appear to be mismatched on some socioeconomic variables and fairly well matched on others. The mothers have approximately the same mean educational level, but the FT families appear to be somewhat larger and to have lower incomes than the NFT families. A considerably larger proportion of the FT children are Black. The two samples perform equally well on the entry level WRAT and PPVT. However, 84 percent of the FT sample attended preschool while only slightly over half of the NFT sample did.

The FT and NFT samples are distributed similarly across the various regions and city sizes. The greatest discrepancy is in the Northeast and North Central large cities, where 43 percent of the FT sample, but only 25 percent of the NFT sample resides. The 1973 and 1974 analytic samples are quite similar in their geographic distributions. The major difference is that last year's NFT sample includes 47 percent residing in small southern cities; this year it includes 26 percent. This results in greater FT/NFT geographic similarity for the current analysis.

There are two major changes in the nature of the traced sample which have significant impact on the interpretation of the High/Scope findings. The criteria used in creating the 1973 analytic sample resulted in the exclusion of all but Black and White children. These criteria were changed



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP COOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	1.668	2.076	1.819	.923	1.054	3.742	-2.366	-.045	.042	-.115
	UNADJ	2.365	3.131	2.462	1.965	1.894	4.701	-2.815	-.033	.203	-.063
S.E. = STANDARD ERROR OF B	ADJ	.411	.463	.306	.638	.446	.518	.658	.017	.209	.177
	UNADJ	.494	.615	.395	.755	.547	.646	.976	.017	.221	.178
F= SIGNIFICANCE STATISTIC	ADJ	** 16.513	** 20.141	** 35.266	2.094	** 5.530	** 52.208	** 12.931	** 7.282	.041	.423
	UNADJ	** 22.885	** 25.898	** 38.835	** 6.767	** 11.976	** 52.945	** 8.321	* 3.593	.843	.124
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.356	.356	.478	.127	.204	.590	-.252	-.282	.021	-.070
	UNADJ	.505	.537	.646	.271	.366	.741	-.299	-.206	.100	-.038
N	FT	320	320	320	320	320	320	320	320	320	320
	NFT	113	113	113	113	113	113	113	113	113	113

1 FOR ALL TESTS,EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT,THE FALL WRAT WAS REPLACED WITH THE FALL PPVT.

** P LESS THAN .05 * P LESS THAN .10

FIGURE V-5 KINDERGARTEN EFFECTS PROFILE FOR KANSAS ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

in creating the traced sample so that a large number of Spanish speaking children (17%=FT; 21%=NFT) were included in the new analyses. In addition, there was an increase of about 3% Chinese children in the NFT group. Obviously, these changes were not distributed equally across all sites or schools; therefore, the impact of potentially low scoring children on both the FT and NFT groups and of some potentially high scoring children on the NFT group could be rather large at some sites and some schools.

5.6.2 Kindergarten 1-Year Effects

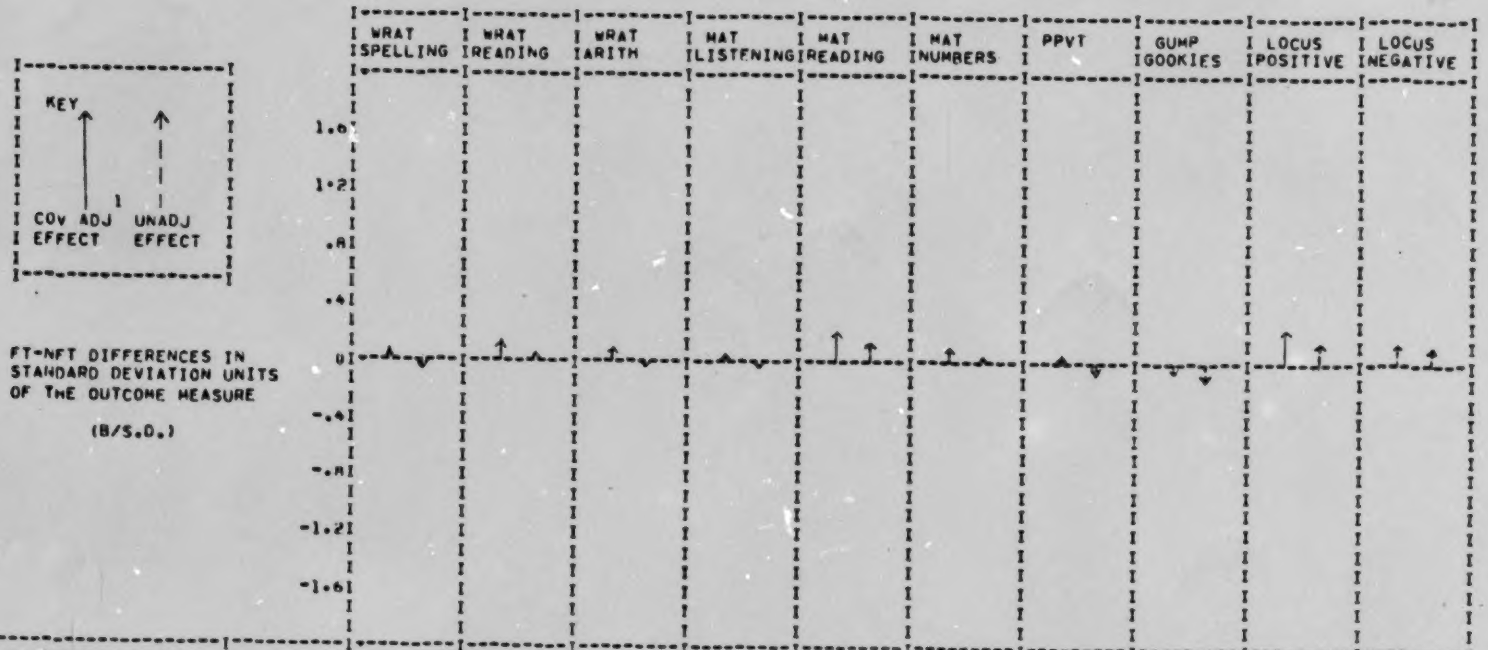
The High/Scope kindergarten FT/NFT contrasts are, as shown in Figure V-6 , similar in pattern to those found in the analyses of the 1973 sample, but they are smaller in magnitude. The very close match in the 1973 sample yielded consistently moderate and positive effects at the school level and somewhat smaller effects in child level analyses. The traced sample, with an equally close match, but with more children in both FT and NFT who could be expected to reduce the differences in mean achievement levels, did show almost identical FT/NFT effects. Although the pattern of consistent FT/NFT equality across outcome measures is clear, it is also clear that the very heterogeneous sample requires a careful examination of these effects at site level for maximum interpretation. These are presented in the next chapter.

5.7 UNIVERSITY OF FLORIDA PARENT EDUCATION MODEL

5.7.1 Sample Description

The FT and NFT samples for the University of Florida this year appear to be similar in some respects and dissimilar in others. The FT and NFT mothers have similar educational levels, but the FT families have much lower incomes than the NFT families. On the other hand, the FT mothers appear to have a much stronger belief in their ability to control external events. There is a considerably higher proportion of Blacks in the FT sample. The two samples have performed equally well on the entry level PPVT and WRAT. A considerably higher proportion of FT children attended some form of preschool.

This Sponsor's FT and NFT samples are located in very different regions and city sizes. Whereas 34 percent of the FT sample and none of the NFT sample is located in the rural West, 47 percent of the NFT sample and none of the FT sample is located in the rural North Central region. A small



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	.194	.834	.377	.439	1.109	.710	.493	-.013	.476	.256
	UNADJ	-.286	.285	-.019	-.186	.722	.009	-.676	-.020	.305	.194
S.E. = STANDARD ERROR OF B	ADJ	.426	.480	.318	.662	.465	.538	.684	.017	.216	.184
	UNADJ	.512	.637	.409	.782	.567	.669	1.011	.018	.229	.185
F= SIGNIFICANCE STATISTIC	ADJ	.207	* 3.020	1.403	.439	** 5.685	1.745	.519	.534	** 4.829	1.933
	UNADJ	.312	.290	.002	.056	1.622	.000	.449	1.250	1.769	1.105
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.041	.143	.099	.061	.214	.112	.052	-.079	.233	.156
	UNADJ	-.061	.049	-.005	-.026	.140	.001	-.072	-.126	.149	.119
N	FT	168	168	168	168	168	168	168	168	168	168
	NFT	145	145	145	145	145	145	145	145	145	145

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

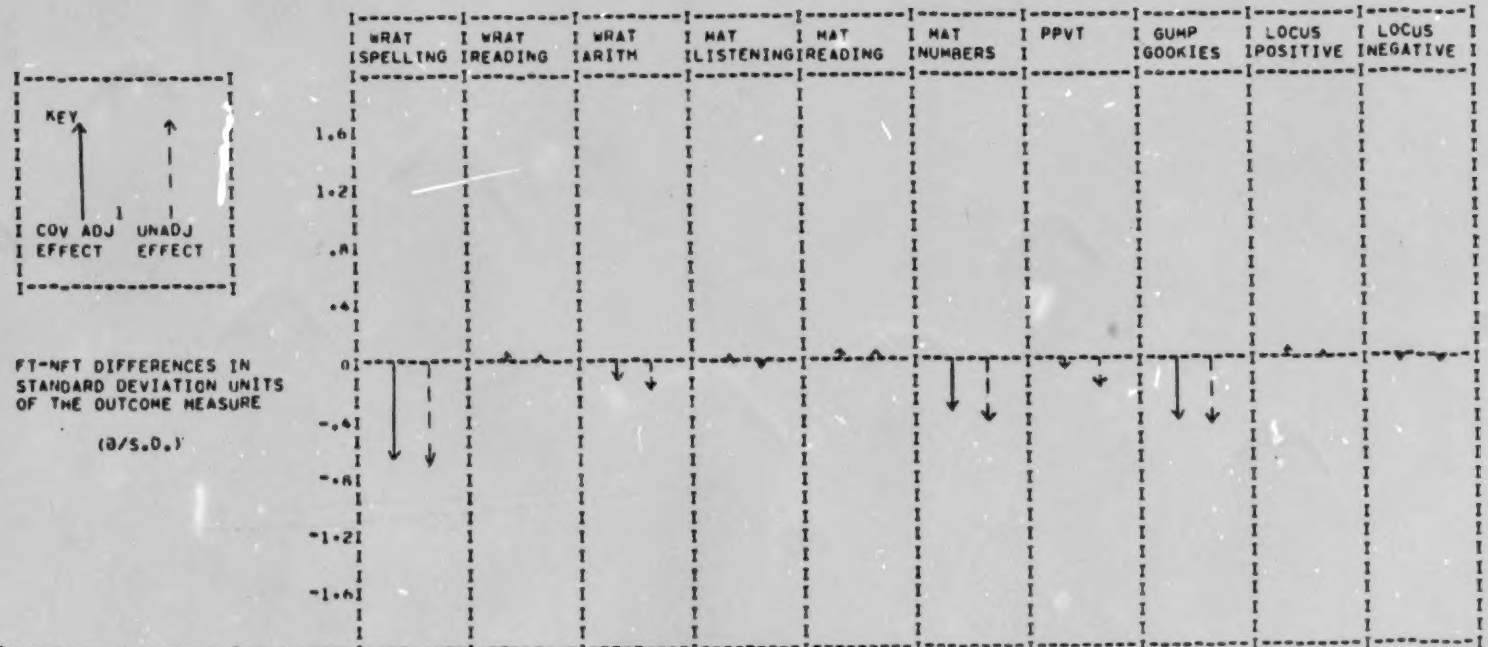
Figure V-6 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

proportion (11%) of Florida's FT sample is also found in the large cities of the Northeast; no NFT sample children reside there. The only FT/NFT geographic similarity is that slightly over half of each sample is located in large Southern cities. Last year's geographic distribution was quite similar to this year's except that there were fewer FT pupils last year in large Southern cities and about 20 percent of the FT sample was in North Central rural areas compared to none for the current FT group.

Last year's FT and NFT samples were also mismatched, with the FT group having lower income. This FT/NFT income differential is significantly increased in the current sample; tracing resulted in a sharp drop in FT mean income, but no change in NFT mean income. Mean level of mother's education in NFT is slightly lower this year. Unlike last year, this year's Florida sample has an FT/NFT mismatch, in favor of the NFT group, on PPVT. Clearly, the application of tracing had very different effects on FT compared to NFT which reflects the fact that these groups are located in different parts of the country. The only substantial change in the sample which has altered the composition of FT and NFT in similar ways is the inclusion of Spanish children. In the current analytic sample, 16% of the FT children and 21% of the NFT children are Spanish. These changes occurred in specific Florida sites, which suggests again that site analyses, which eliminate geographical differences in FT and NFT, are critical for an understanding of this Sponsor's effects.

5.7.2 Kindergarten 1-Year Effects

The kindergarten FT/NFT effects for Florida, shown in Figure V-7 indicate no difference between FT and NFT on some measures, but a substantial NFT-favoring difference on the WRAT Spelling subtest, on the MAT Numbers subtest, and on the Gumpgookies. This pattern is dramatically different from last year's kindergarten analysis where Florida showed substantial FT-favoring differences on most outcome measures. The differences between the two analyses can be accounted for in terms of the sample differences across analyses. However, the differences between these results suggest that the Florida model interacts or operates differently with different groups of children. The variation in the pattern of kindergarten FT/NFT differences for Florida's sites, shown in Appendix C, Figures C-26 through C-29, substantiates this hypothesis and further suggests that the factor with which the Florida program interacts is a site-specific variable. Two of Florida's sites show



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-3.160	.444	-.568	.028	.342	-2.322	-.581	-.071	.159	-.056
	UNADJ	-3.445	.221	-.774	-.289	.232	-2.769	-1.965	-.074	.013	-.110
S.E. = STANDARD ERROR OF B	ADJ	.523	.589	.390	.812	.570	.659	.839	.021	.2[illegible]	.225
	UNADJ	.627	.781	.502	.959	.695	.820	1.239	.022	.28[illegible]	.226
F= SIGNIFICANCE STATISTIC	ADJ	** 36.559	.567	2.124	.001	.360	** 12.407	.480	** 11.108	.359	.061
	UNADJ	** 30.146	.080	2.380	.091	.112	** 11.402	2.518	** 11.503	.002	.234
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.8[illegible]	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.674	.076	-.149	.004	.066	-.366	-.062	-.443	.078	-.034
	UNADJ	-.735	.038	-.203	-.040	.045	-.437	-.209	-.467	.007	-.067
N	FT	173	173	173	173	173	173	173	173	173	173
	NFT	74	74	74	74	74	74	74	74	74	74

[1] FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

Figure V-7 KINDERGARTEN EFFECTS PROFILE FOR FLORIDA ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

FT-favoring effects on most outcomes, and two other sites show the opposite. The possibility of an interaction of the model and some site effects will be explored in greater detail in the next chapter.

5.8 THE EDUCATIONAL DEVELOPMENT CENTER (EDC)

5.8.1 Sample Description

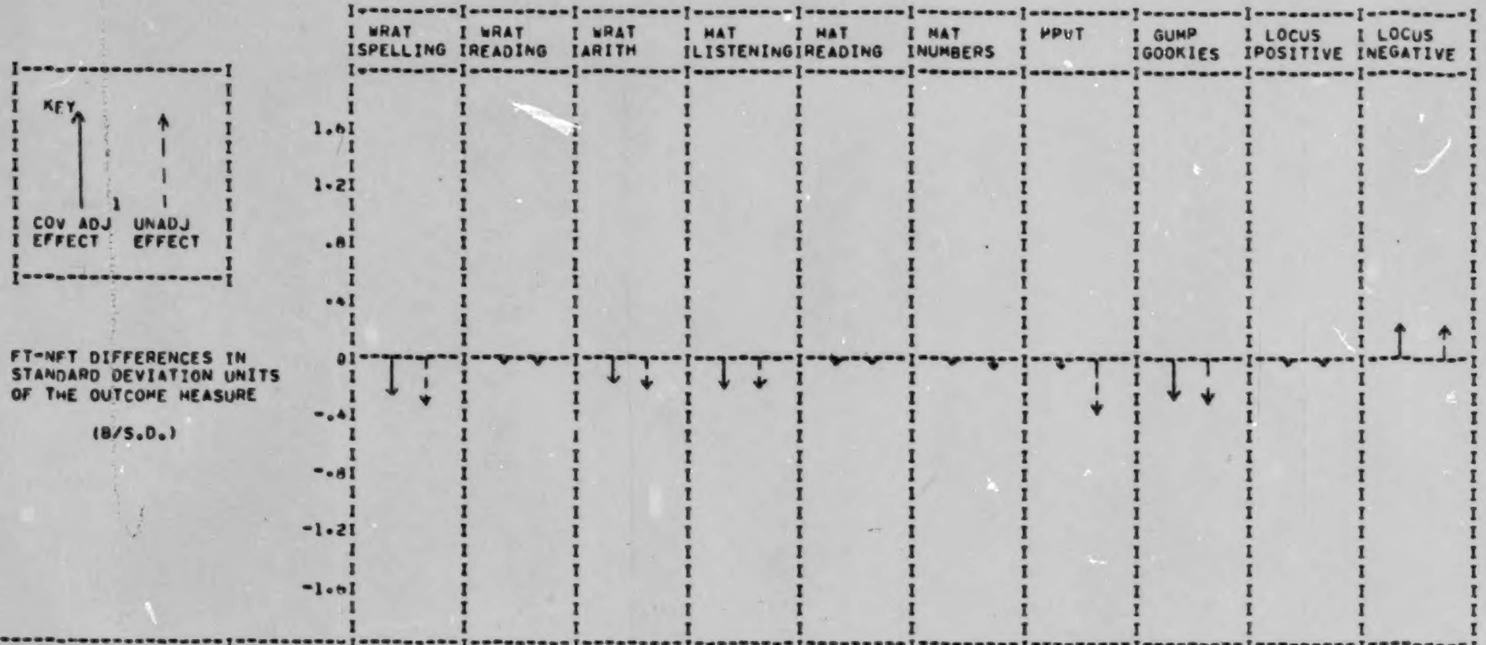
The FT and NFT samples in this year's analysis for EDC are quite similar socioeconomically. The major difference between the two groups is that there are more Blacks in the FT sample. Although the groups did equally well on the entering WRAT, the FT groups did less well on the PPVT. A considerably larger proportion of the FT group attended preschool and the FT group is also more than one month younger than the NFT sample.

All pupils in this year's EDC samples reside in the Northeast; slightly less than one-third of both FT and NFT groups resides in medium-sized cities there. The remainder of both groups is in large and small cities. However, 52 percent of the FT sample and 7 percent of the NFT sample reside in small cities. Last year's sample is quite similar to this year's sample except that there were fewer FT children last year in the large and medium cities of the Northeast and more in Northeastern small cities. Last year's FT sample also had 11 percent of the children in large Southern cities; none are located there in the current analysis. The NFT sample last year had more children in Northeastern small cities and fewer in Northeastern medium-sized cities than the current NFT sample.

Last year the FT and NFT samples were quite mismatched with the FT sample having lower incomes and less educated mothers. Hence, the fairly close FT/NFT match found in the current sample should give us a better chance at estimating EDC's effect.

5.8.2 Kindergarten 1-Year Effects

The kindergarten FT/NFT differences for EDC, shown in Figure V-8, indicate few differences between the groups at the end of kindergarten, duplicating last year's results. The kindergarten effects for EDC's sites, shown in Appendix C, Figures C-30 to C-32, however, indicate a high degree of variability. Site analyses, presented in the next chapter, are critical for understanding this Sponsor's effects.



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.336	-.043	-.736	-1.477	-.021	-.165	-.616	-.045	-.022	.383
	UNADJ	-1.475	-.172	-.846	-1.643	-.110	-.368	-3.410	-.047	-.070	.364
S.E. = STANDARD ERROR OF B	ADJ	.490	.552	.365	.761	.534	.618	.788	.020	.249	.211
	UNADJ	.590	.735	.472	.902	.654	.772	1.166	.021	.265	.213
F= SIGNIFICANCE STATISTIC	ADJ	** 7.443	.006	** 4.060	* 3.766	.002	.071	.610	** 5.115	.008	* 3.287
	UNADJ	** 6.245	.055	* 3.216	* 3.317	.028	.227	** 8.562	** 5.179	.070	* 2.914
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.285	-.007	-.193	-.204	-.004	-.026	-.065	-.282	-.011	.234
	UNADJ	-.315	-.029	-.222	-.227	-.021	-.058	-.363	-.295	-.034	.222
N	FT	150	150	150	150	150	150	150	150	150	150
	NFT	96	96	96	96	96	96	96	96	96	96

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

Figure V-8 KINDERGARTEN EFFECTS PROFILE FOR EDC ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

5.9 UNIVERSITY OF PITTSBURGH, PRIMARY EDUCATION PROJECT

5.9.1 Sample Description

Along with Bank Street, this Sponsor suffered extreme losses in sample size and characteristics due to the two year tracing procedure applied to the 1973 kindergarten sample. The loss of NFT children was particularly great. Here only 38% of the 1973 analytic sample was retained, compared to 67% of the national NFT analytic sample. The FT traced sample retained only 59% whereas the national FT retention rate was 75%. These losses were not evenly distributed over the locations in which Pittsburgh was administering its programs. First, there was a complete loss of the large city site as a result of the tracing; 69 FT and 25 NFT children were lost here. Next, about one-third of the FT and one-third of the NFT children were lost from the small city site. However, in the rural area, virtually all of the FT children were retained (110 out of 113), while two-thirds of the NFT children were lost (58 out of 91). Clearly, very different events were occurring at these different locations, and the processes which led to testing and tracing children were different for the FT than for the NFT groups. Changes in the characteristics of children over the two samples demonstrate this point.

The major change in the FT sample when it was subjected to the tracing procedure was the sharp increase in the mean family income. Since the largest loss of FT children occurred in the loss of the large city site, it is clear that these children came from relatively low income families in the Pittsburgh distribution of family income. Associated with this is a significant increase in the mean PPVT score of the FT traced group. Small, non-significant increases in the WRAT entry scores and in mother's education are also apparent in the traced group. Thus, the FT traced group lacks the low income, lower achieving large city children who were included in the 1973 sample.

The major change in the traced NFT group, other than size, is a very sharp increase in WRAT reading entry score. There are small, non-significant increases in the other entry level scores despite the fact that there is no change in the SES of the traced group compared to the 1973 group. The NFT children who were lost come from the same SES as those retained, but they scored much lower on the reading test at time of entry into the program. These children came primarily from the rural areas involved with the Pittsburgh program. Thus, the traced NFT sample lacks the upper income, lower achieving, rural children who were included in the 1973 sample.

The result of the divergent tracing processes is an FT/NFT traced sample which is very much more matched than the sample last year. The groups are predominantly White, have high levels of mother's education and family income, and are among the highest performing children on entry into the program in the national group of both FT and NFT children. The only important difference between the FT and NFT traced groups is the higher percentage of FT children who attended preschool (86% vs. 31%).

The current sample is located in two sites; 70% of the FT and 55% of the NFT sample are located in a North Central rural area. The remainder of both groups is located in a small city in the Northeast. These are hardly representative of the Pittsburgh involvement in Follow Through and ought not to be used as a basis for generalization of findings to the overall Pittsburgh effects which clearly have not been included in these analyses.

5.9.2 Kindergarten 1-Year Effects

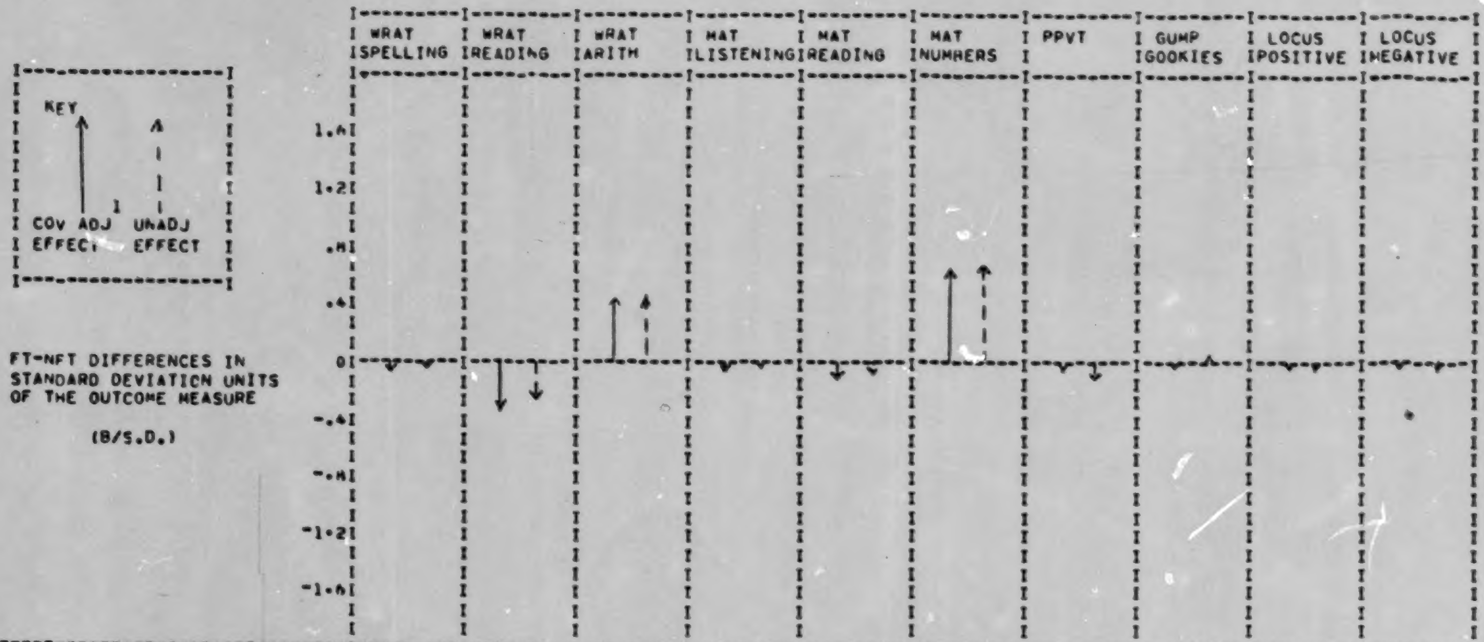
The 1973 analytic sample showed positive effects of FT in all outcome measures (See Figure V-9). The traced sample shows such positive effects only in MAT Numbers and WRAT Arithmetic. NFT shows an advantage in the reading domain. Clearly, the sharp increase in the entry characteristics of the NFT group, particularly in reading, has contributed strongly to this change in the effects. This is true despite the increase in SES of the traced sample which is sometimes predictive of higher performance levels. Once again, the effects by site, when some of these sample changes are sorted out, will show more about the Pittsburgh effects than can be seen at the overall Sponsor level.

5.10 SOUTHWEST EDUCATION DEVELOPMENT LABORATORY (SEDL)

5.10.1 Sample Description

SEDL's FT and NFT samples are quite similar this year socioeconomically. The FT sample scored lower than the NFT sample on the entering PPVT and WRAT subtests. Although the FT children are younger, more of them attended preschool than did their NFT counterparts.

The current FT/NFT samples, both located primarily in rural areas, come from different regions of the country. An overwhelming majority (94%) of the FT sample resides in the rural South, while less than one-quarter of the NFT sample does. Most of the NFT sample (70%) resides in the rural West; none of the FT sample does. The remainder of each sample (about 7%) resides in large Northeastern cities.



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP GOOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.334	-1.881	1.657	-.425	-.570	4.095	-.569	-.001	-.005	-.050
	UNADJ	-.220	-1.685	1.774	-.242	-.392	4.234	-1.071	.001	-.005	-.049
S.E. = STANDARD ERROR OF B	ADJ	.571	.643	.426	.887	.623	.720	.916	.023	.290	.246
	UNADJ	.688	.857	.550	1.051	.762	.899	1.358	.024	.308	.248
F= SIGNIFICANCE STATISTIC	ADJ	.343	** 8.560	** 15.145	.230	.839	** 32.347	.386	.002	.000	.040
	UNADJ	.102	** 3.868	** 10.403	.053	.264	** 22.164	.622	.003	.000	.039
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SFE GRAPH)	ADJ	-.071	-.323	.435	-.059	-.110	.646	-.060	-.006	-.002	-.030
	UNADJ	-.047	-.289	.466	-.033	-.076	.668	-.114	.008	-.002	-.030
N	FT	153	153	153	153	153	153	153	153	153	153
	NFT	60	60	60	60	60	60	60	60	60	60

1 FOR ALL TESTS, EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

Figure V-9 KINDERGARTEN EFFECTS PROFILE FOR PITTSBURGH ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

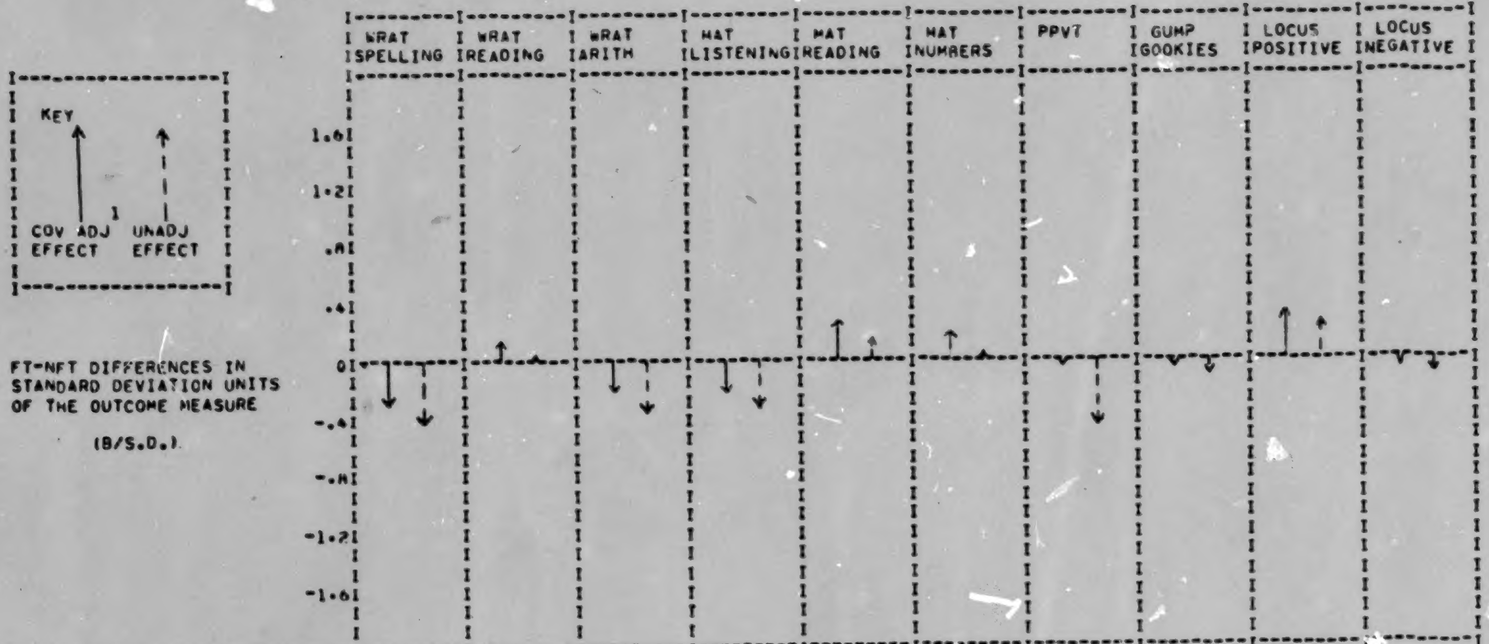
Although the samples were also primarily rural last year, unlike this year the NFT rural sample was split evenly across Western and Southern sites. Last year the FT sample was also represented by a rural western site which is not included in the current analysis.

SEDL's FT/NFT samples were more closely matched last year than they are this year. Much of the difference between samples may be related to the fact that this year's FT group is 34% Spanish-speaking while the NFT group is only 18% Spanish-speaking. This dramatic alteration in the ethnic composition of the sample may account for the absence of SES changes across samples with the substantial drop in entry test scores. The first language of many of these children is Spanish, which may produce more difficulty in early testing than later testing.

5.10.2 Kindergarten 1-Year Effects

The kindergarten FT/NFT differences for SEDL, shown in Figure V-10, indicate few differences between FT and NFT on kindergarten outcome measures. This profile of differences represents a substantial change from last year's analysis where SEDL shows FT-favoring effects on many outcome measures and particularly large effects in reading. The differences between the two analyses suggest that the SEDL program is more effective in kindergarten for English-speaking children and less effective for children with Spanish as a first language.

SEDL's kindergarten site effects, shown in Appendix C, Figures C-35 through C-37 show a high degree of variability. Although each site shows some FT favoring effects, they are not consistent across outcome measures or domains. SEDL does not produce a consistent pattern of effects in kindergarten, and the model may be more or less effective in kindergarten, depending on the type of children involved.



		WRAT SPELLING	WRAT READING	WRAT ARITH	MAT LISTENING	MAT READING	MAT NUMBERS	PPVT	GUMP COOKIES	LOCUS POSITIVE	LOCUS NEGATIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.450	.865	-.892	-1.628	1.369	1.157	-.301	-.010	.681	-.112
	UNADJ	-2.060	.036	-1.430	-2.486	.739	.299	-4.417	-.020	.514	-.170
S.E. = STANDARD ERROR OF B	ADJ	.700	.789	.523	1.088	.764	.883	1.126	.028	.356	.302
	UNADJ	.844	1.051	.675	1.290	.935	1.103	1.666	.029	.378	.305
F= SIGNIFICANCE STATISTIC	ADJ	** 4.288	1.177	* 2.918	2.239	* 3.210	1.716	.071	.125	* 3.665	.137
	UNADJ	** 5.955	.091	** 4.492	* 3.715	.625	.074	** 7.024	.468	1.844	.312
SD=RAW SCORE STANDARD DEVIATION		4.685	5.829	3.809	7.247	5.171	6.340	9.403	.159	2.044	1.636
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.310	.148	-.234	-.225	.265	.183	-.032	-.063	.333	-.068
	UNADJ	-.440	.006	-.375	-.343	.143	.047	-.470	-.127	.251	-.104
N	FT	67	67	67	67	67	67	67	67	67	67
	NFT	50	50	50	50	50	50	50	50	50	50

1 FOR ALL TESTS EXCEPT PPVT, THE ADJUSTING COVARIATES ARE FALL WRAT, INCOME, AND MOTHERS EDUCATION. FOR THE PPVT, THE FALL WRAT WAS REPLACED WITH THE FALL PPVT. ** P LESS THAN .05 * P LESS THAN .10

Figure V-10 KINDERGARTEN EFFECTS PROFILE FOR SEDL ON CHILD LEVEL ANCOVA ANALYTIC SAMPLE

5.11 CONCLUSIONS

In general, sample differences as a result of the tracing are very variable from one Sponsor to the next. FT/NFT differences on entry or demographic characteristics went up in some cases, down in others, and remained the same in others despite significant changes in the number of children included in the two samples.

In every case but one, changes in the kindergarten effects from the 1973 study to the 1974 study follow directly and logically from the nature of the changes in the sample. Thus, an increase in the mismatch produced a larger difference in the effects in the direction of the higher performing or higher SES group. When low achieving children were systematically lost from a particular group, the effects were increased for that group. When non-English speaking children were added to the sample, effects tended to be diminished.

The one instance in which the sample differences did not appear to be associated with changes in effects in an expected direction, is the case of Far West. Here the sample changed toward more economically disadvantaged children with lower entry scores. The net effect of this change was an increase in the FT factoring effect. This is not in the expected direction, so an ad hoc explanation was asserted: Far West is particularly effective with low achieving children. This explanation shall be examined later on with two year data but the changes in sample and the associated effects at least appear to be rationally related to each other. In all other cases we feel confident that the effects found last year have been replicated this year since all changes in the effects can be rationally explained on the grounds of sample changes. Given these findings, we can go on to the next step in the national evaluation: the two-year effects.

CHAPTER VI

TWO-YEAR EFFECTS

6.0 INTRODUCTION

The purpose of this chapter is to extend the examination of FT/NFT contrasts to a group of children who have experienced two continuous years of Follow Through with the same Sponsors. The two year traced sample, described in Appendix A, is thus the focus of this chapter.

The goal of the analyses reported here is to identify the size of the contrasts between FT and NFT at the end of first grade. Note that this does not mean the effects of the first grade experiences alone. Rather, it refers to the effects of two continuous years covering kindergarten and first grade and culminating in test performance at the end of first grade. Since this analysis must be applied to each Sponsor, the sample required includes only those children who remained in the program for these two continuous years with the same Sponsor. Such children for whom full sets of data are available constitute the analytic groups discussed in this chapter.

Because a two-year effect is the goal, the most appropriate means of adjusting for initial differences between FT and NFT involves the use of pre-kindergarten measures as covariables. This is the logic and procedure used in accomplishing the one-year kindergarten analyses reported in the previous chapter. Consequently, precisely the same analytic model was applied to precisely the same group of children to accomplish the two-year effects analyses as in the case of the one-year effects analyses, except that the dependent measures used in the two-year study included those administered at the end of first grade, whereas the dependent measures used in the one-year study were, of course, those administered at the end of kindergarten. These batteries of measures differed in at least two important respects. First, the achievement measures used at each grade level were different instruments so that somewhat different content and subscales are found in these batteries. The second difference is that whereas the kindergarten battery includes affective and motivational measures, the first grade battery does not.

Although all of the effects are reported in standardized form to enhance comparison across measures within the same Sponsor,[1] it is not the purpose of this chapter to compare the effects produced in one year with the effects produced in two years of continuous FT experience. The reader interested in the longitudinal relationships between these effects is referred to Monograph I in Volume IIB of this report.

An additional aspect of this chapter which differentiates it from the previous chapter is the site analyses. Here, analyses of the traced sample organized by individual site for each Sponsor will be reported along with analyses of all children summed over sites for each Sponsor. For each Sponsor, therefore, a series of analyses will be reported in which the entire set of FT children for that Sponsor who meet the inclusion criteria for participation in the analysis is contrasted with the entire set of NFT children who likewise meet the inclusion criteria. Then a series of analyses will be reported in which the FT children meeting the inclusion criteria within a given site will be contrasted with the NFT children meeting the criteria in the same site. It should be noted that the sum of the individual site analyses does not always equal the overall Sponsor analysis, because the numbers of children and of sites do not always match in these two ways of organizing the data. The reason for this is that one of the inclusion criteria for the site level analyses is that a site must have at least 16 children represented in each of the experimental categories of FT and NFT. This criterion was not, of course, applied to the overall analysis. A site which contained fewer than 16 children in one of the experimental categories would therefore be removed from the site analyses, whereas the children from that site would appear in the overall analysis of that Sponsor. These differences show up occasionally in the form of a discrepancy between the overall results for a Sponsor and the apparent average of the several sites involved in the site analyses. Interpretation of such findings are not materially affected by these discrepencies.

It is appropriate to indicate here the reasons for the addition of the site analyses to the present chapter. There are three reasons for these additional analyses:

---

[1] Comparisons across Sponsors are very tenuous and should not be attempted without careful consideration given to the many extenuating circumstances discussed in Volume I of this series (Cline et. al., 1974).

The first reason is that there is a good deal of anecdotal evidence that very divergent socioeducational events are occurring across sites for a given Sponsor. These include the issues of implementation (discussed in Monograph III, Volume IB of this series) which appear to be operating in somewhat different ways in different sites for some Sponsors. The anecdotal evidence suggests that sites are very different in respect to the administration and operation of FT programs and in the characteristics of the comparison NFT groups. Combining sites means combining diverse events, which tends to obscure important processes.

The second reason for the site analyses is that the discrepencies between some of the kindergarten effects found with the traced two-year sample and those found for the one-year kindergarten sample (discussed in the previous chapter) appear in some cases to be related to site differences in the two samples. That is, in some cases the loss of children in the traced group was not evenly distributed over all sites but was concentrated in some sites, and the change in effects between the two sets of analyses appears to be associated with the loss of whole sites and/or disproportionate numbers of children from a few sites. Looking at individual sites is thus justified on these grounds as well as on anecdotal grounds.

The third reason for the site analyses is that a substantial amount of heterogeneity of regression occurs in the overall Sponsor analyses (see Appendix B for details). This means that for a given Sponsor on a given outcome measure, the covariable (i.e., the pretest) performs its adjusting function in one way for the FT group and in another way for the NFT group, and it therefore removes non-comparable portions of the variance from these two groups. Adjustments for differences in pretest may be spuriously high in one group and spuriously low in the other when heterogeneous regressions are ignored, and comparisons between the groups can be gravely misleading.

The critical factor here is that heterogeneity of regression may be produced in overall Sponsor analyses precisely because the sites for a given Sponsor are combined in that Sponsor's analyses. The overall heterogeneity may occur because each site shows its own distinctive relationships between pretest and outcomes in terms of levels and regressions. Just as the anecdotal evidence suggests that sites often contain very different educational dynamics, the heterogeneity findings suggest that sites are quite different and should often not be combined in analyses. The demonstration that this is most likely the case is found in the elimination of heterogeneity when analyses are performed on each site. When taken one at a time, the sites are homogeneous and can be analyzed in the ANCOVA mode.

Site analyses also represent a form of replication of findings to which we have referred earlier as an aspect of the stability of Sponsor effects from place to place. It would be difficult to make any policy decisions about any program if one did not know how likely the observed effects would be to emerge in new sites. The sites in this analysis were not selected by any sampling plan, and so there is no formal statistical justification for any generalization of findings. Nevertheless, it is reasonable to look for stability or instability of findings over several different locations, simply to get a first tentative estimate of the sensitivity of the various FT programs to different conditions and locations of operations. For this reason as well, we report individual site analyses.

The format for presentation of findings is identical to that used in the previous chapter.

6.1 FAR WEST LABORATORY, RESPONSIVE EDUCATIONAL PROGRAM

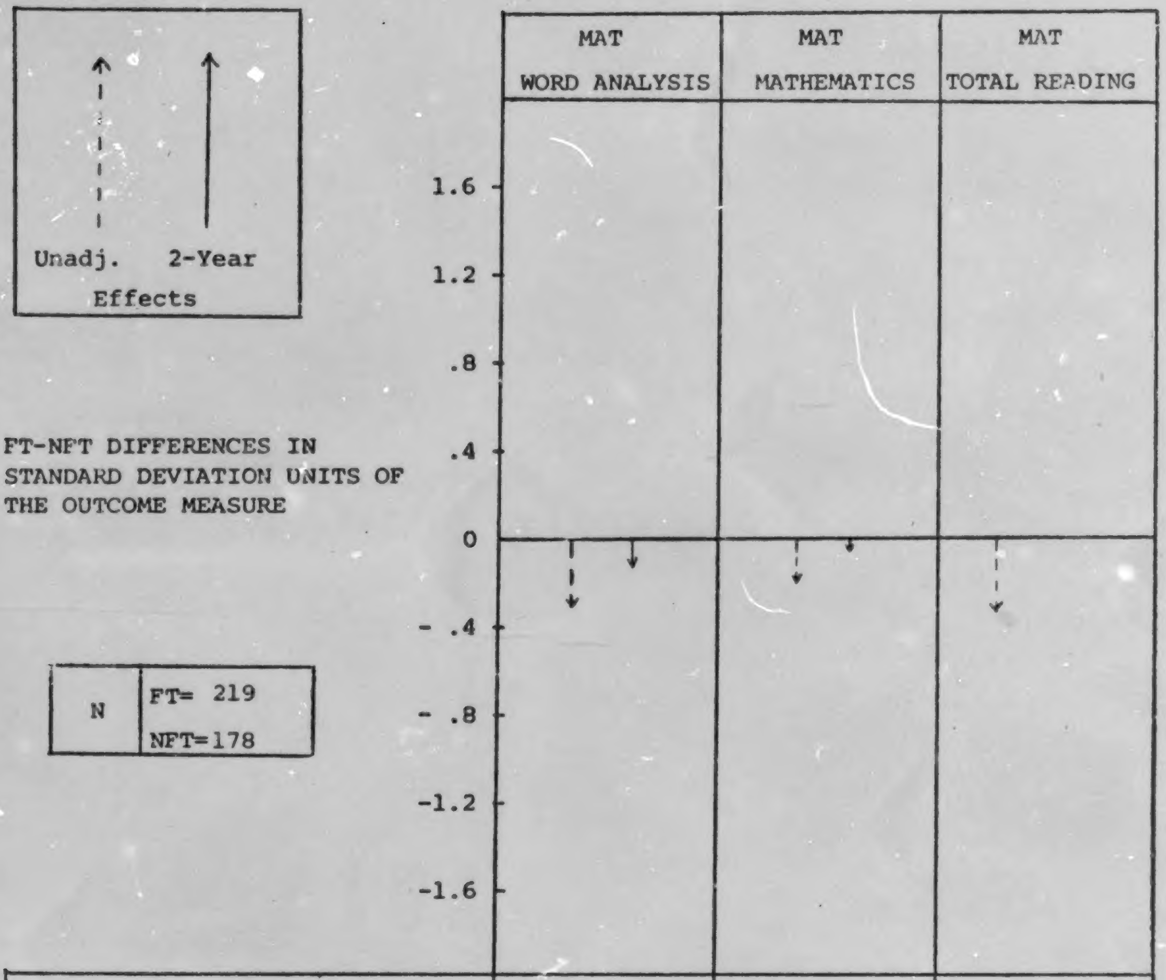
The Sponsor describes the Responsive Educational Program as one that addresses the variety of learning needs of children while responding to the variety of needs of parents, schools, and communities. Responsive teaching focuses on a child's language development and problem solving skills to produce an independent learner. An independent learner has the skills to solve a variety of problems and the confidence to attempt them. The responsive classroom encourages exploration and discovery learning; the child's life experiences are integrated into classroom activities.

6.1.1 Overall 2-Year Effects

The Far West FT/NFT differences at the end of first grade, shown in Figure VI-1, indicate little difference between the groups on Word Analysis and Mathematics. Only the unadjusted score on Total Reading is shown; an adjusted difference could not be computed because of heterogeneity of the relationship between the Total Reading scores and entry WRAT scores for FT and NFT. Figure VI-2 shows the initial advantage of the NFT group on entry WRAT and its posttest advantage on the Total Reading score. The difference in the slopes of the two regression lines suggests that higher entry level NFT children gained relatively more than higher entry level FT children. In any case, a substantial part of the NFT advantage on the posttest is attributable to pretest differences. Taking into consideration the FT disadvantage on pretests plus the disadvantage of the FT children on SES, it is likely that the substantial posttest difference on Total Reading (.35 standard deviations) represents background and entry differences and not differences attributable to kindergarten and first grade experience. The FT children thus appear to be maintaining their status with respect to the NFT children in reading.

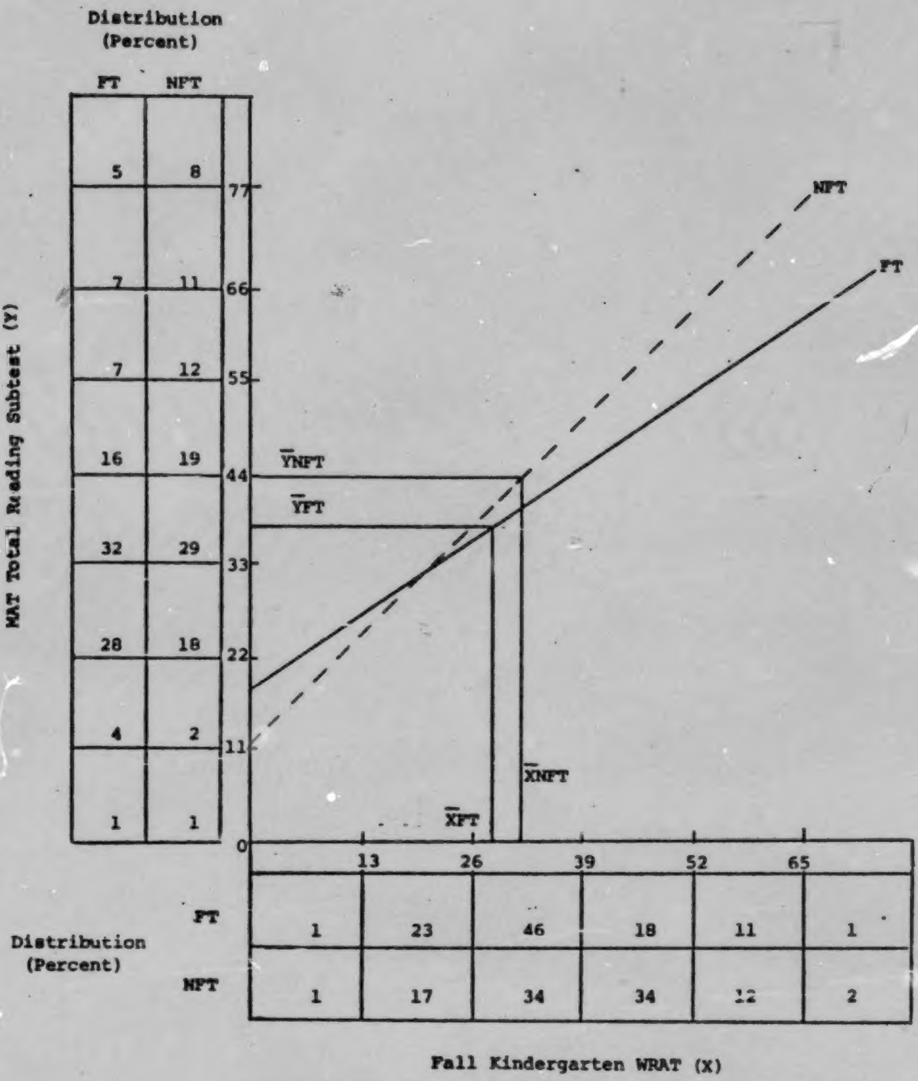
The FT group is representative of a disadvantaged group; it is lower on SES measures, has lower entry level achievement scores, and has a substantially larger percentage (80% vs. 27%) of children with Head Start or other preschool experience than the NFT group. Since traditional findings (e.g. Coleman, 1968, and Beller, 1973) suggest that the gap between advantaged and disadvantaged children widens throughout the school years, the fact of the FT group's maintenance of entry status over the two year period covering kindergarten and first grade may represent a favorable treatment effect.

FIGURE VI-1 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FAR WEST



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .083	- 2.784	- 6.198
	2-Year	- .037	- .781	
S.E. = Standard Error of B	Unadjusted	.026	1.178	1.661
	2-Year	.020	.953	
F = Significance Statistic	Unadjusted	10.239*	5.588*	13.923*
	2-Year	3.270	.672	
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .310	- .213	- .354
	2-Year	- .138	- .060	

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education.
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05



	FT	NFT
$\bar{X}$	28.70	31.93
$\bar{Y}$	37.79	43.98
b	.83	1.02
SE (b)	.08	.08
n	219	178

Figure VI-2

Heterogeneous Regressions:
Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT
Sponsor: Far West
Outcome: MAT Total Reading Subtest

This interpretation is tentative as well as open to criticism. The degree of difference between FT and NFT on SES may not be sufficiently large to produce a divergence in the level of achievement of the two groups which is discernable at the end of the first grade. The validity of this interpretation cannot, however, be determined until data from subsequent grades are analyzed and the continuity of the trend determined.

6.1.2 Site Description

All five sites used for the analysis of the Far West Program are metropolitan communities of 100,000 persons or more. However, they are scattered geographically and the children in the programs differ greatly in ethnic background, language, SES, and initial achievement. The pupil and family background characteristics for each of the five sites included in these analyses are shown in Tables VI-1 to VI-5.

The first site, a medium-sized North Central city (NC1, Table VI-1), differs from the first not only in size and region, but in at least two of the characteristics of the children served. Here the FT/NFT children are predominantly White (FT = 92%; NFT = 93%), and the two groups differ substantially on family income and family structure. They are, nevertheless, well matched at kindergarten entry level.

At the second site, a large North Central city (NC2, Table VI-2), both FT and NFT groups are virtually all Black and have similar SES on all indicators. FT children scored lower at entry to kindergarten than their NFT counterparts (FT = 36.63; NFT = 42.26) on the PPVT; the groups are well matched on entry level WRAT.

The third site (NE, Table VI-3) is a large Northeastern city, serving predominantly Black (FT = 100%; NFT = 71%) and lower SES children. The two groups are similar in kindergarten entry level WRAT, but the FT children scored somewhat lower on the PPVT. The FT group in the site appears to be slightly disadvantaged relative to NFT.

At the fourth site, a medium-sized Western city (W1, Table VI-4), approximately two-thirds of the children are Black (FT = 60%; NFT = 70%) and one-third Spanish-surnamed (FT = 35%; NFT = 24%). The FT group has more

TABLE VI-1

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR FAR WEST SITE NC1

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.53	.82	83	5.72	.74	65
Father's Education	5.58	.99	55	5.75	1.05	60
Income	10.96	7.04	78	15.16	6.31	68
Occupation	15.26	2.07	81	15.22	2.12	64
Family Structure						
Mother and Father Present	63%	-	53	82%	-	56
Mother, not Father, Present	33	-	26	11	-	7
Other	4	-	4	7	-	5
TOTAL			83			68
Entry Level						
WRAT	32.78	11.06	73	34.31	10.37	51
PPVT	52.20	7.40	73	50.84	8.26	51
Preschool Experience						
Attended Headstart or other Preschool	50%	-	46	23%	-	16
TOTAL			93			69
Ethnicity						
Black	4%	-	3	1%	-	1
White	92	-	67	93	-	47
Spanish	0	-	0	0	-	0
Other	4	-	3	6	-	3
TOTAL			73			51

TABLE VI-2

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR FAR WEST SITE NC2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.32	1.04	34	4.83	1.44	36
Father's Education	missing			4.38	1.50	13
Income	6.92	4.76	36	6.76	6.91	37
Occupation	16.06	1.95	34	16.17	1.52	36
Family Structure						
Mother and Father Present	22%	-	8	26%	-	10
Mother, not Father, Present	72	-	26	58	-	23
Other	6	-	2	16	-	6
TOTAL			36			39
Entry Level						
WRAT	27.46	9.90	41	28.80	10.31	35
PPVT	36.63	9.96	41	42.26	11.84	35
Preschool Experience						
Attended Headstart or other Preschool	76%	-	16	33%	-	8
TOTAL			21			24
Ethnicity						
Black	98%	-	40	100%	-	35
White	0	-	0	0	-	0
Spanish	2	-	1	0	-	0
Other	0	-	0	0	-	0
TOTAL			41			35

TABLE VI-3

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR FAR WEST SITE NE

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.49	.80	53	5.46	.86	26
Father's Education	5.59	.91	22	5.25	1.48	12
Income	7.78	6.08	49	8.04	6.33	25
Occupation	15.80	1.58	44	16.82	1.10	22
Family Structure						
Mother and Father Present	40%	-	23	50%	-	13
Mother, not Father, Present	55	-	31	50	-	13
Other	5	=	3	0	-	0
TOTAL			57			26
Entry Level						
WRAT	25.51	10.38	65	27.47	11.40	17
PPVT	39.80	11.64	65	45.53	13.61	17
Preschool Experience						
Attended Headstart or other Preschool	90%	-	80	84%	-	21
TOTAL			89			25
Ethnicity						
Black	100%	-	65	71%	-	12
White	0	-	0	29	-	5
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			65			17

TABLE VI-4

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR FAR WEST SITE W1

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.25	1.19	40	5.32	1.17	31
Father's Education	4.65	1.46	20	6.00	1.41	12
Income	6.86	6.04	42	5.88	6.23	32
Occupation	16.51	1.66	37	15.59	2.31	27
Family Structure						
Mother and Father Present	45%	-	19	28%	-	9
Mother, not Father, Present	45	-	19	60	-	19
Other	10	-	4	12	-	4
TOTAL			42			32
Entry Level						
WRAT	22.15	10.16	33	26.82	12.22	27
PPVT	44.09	9.75	33	45.37	11.94	27
Preschool Experience						
Attended Headstart or other Preschool	88%	-	35	77%	-	13
TOTAL			40			17
Ethnicity						
Black	60%	-	20	70%	-	19
White	5	-	2	6	-	2
Spanish	35	-	11	24	-	6
Other	0	-	0	0	-	0
TOTAL			33			27

intact families and slightly higher family incomes, but is lower on level of father's education. The FT children enter kindergarten well behind the NFT children on the WRAT (FT = 22.15; NFT = 26.82), though not on the PPVT (FT = 44.09; NFT = 45.37). The inconsistency of the differences between FT and NFT in this site are difficult to assess. However, the somewhat higher income in FT may be related to the higher percentage of intact families thus giving the FT group more wage earners per family. The difference between the groups on father's education may be revealing. The NFT group, at least for the father, for whom data are available, averaged a completed high school education whereas the FT fathers averaged less than an 8th grade education. This substantial difference in education level paired with the entry achievement difference suggests that the FT group is somewhat disadvantaged relative to NFT.

At the last site, another medium-sized Western city (W2, Table VI-5), the FT/NFT children are primarily white (FT = 62%; NFT 77%), although the groups contain Spanish-surnamed children (FT = 22%; NFT = 17%) as well. In this community, the FT children have a substantially lower SES and somewhat fewer intact families. In addition, the FT children are slightly behind at kindergarten entry level,reflecting their general entry disadvantage.

6.1.3 By Site 2-Year Effects

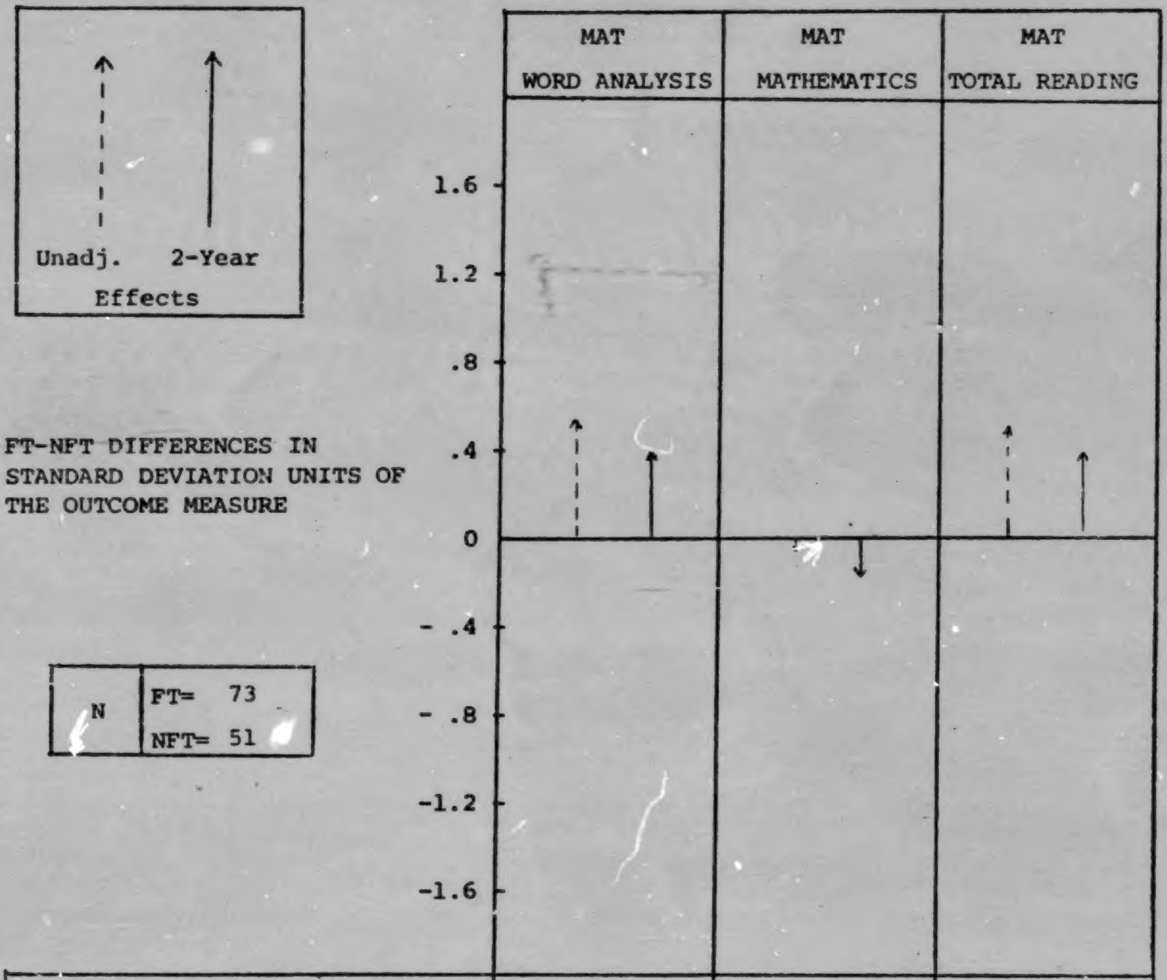
The FT/NFT differences for Far West sites at the end of first grade shown in Figures VL-3 through VI-7 appear to be highly variable. However, the pattern of favorable FT differences across sites corresponds fairly closely to the pattern of mismatch between FT and NFT in the sites. In sites NC1 (Figure VI-3) and NC2 (Figure VI-4) where the Far West FT groups are closely matched with NFT, the FT group shows greater gain over the two years on most outcomes. In both sites the FT groups are substantially ahead of NFT on Word Analysis and Total Reading. In NC2 the FT group is also ahead in Mathematics while site NC1 shows no difference between FT and NFT on this outcome.

TABLE VI-5

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR FAR WEST SITE W2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.30	1.15	44	5.69	1.02	52
Father's Education	5.52	1.26	25	6.03	1.22	39
Income	7.95	6.12	43	13.47	7.03	51
Occupation	15.67	1.79	42	14.52	2.12	50
Family Structure						
Mother and Father Present	52%	-	23	69%	-	36
Mother, not Father, Present	44	-	19	25	-	13
Other	4	-	2	6	-	3
TOTAL			44			52
Entry Level						
WRAT	30.92	11.60	51	33.86	12.43	57
PPVT	49.55	9.91	51	51.75	9.00	57
Preschool Experience						
Attended Headstart or other Preschool	73%	-	37	10%	-	6
TOTAL			51			58
Ethnicity						
Black	14%	-	7	3%	-	2
White	62	-	32	77	-	43
Spanish	22	-	11	17	-	10
Other	2	-	1	3	-	2
TOTAL			51			57

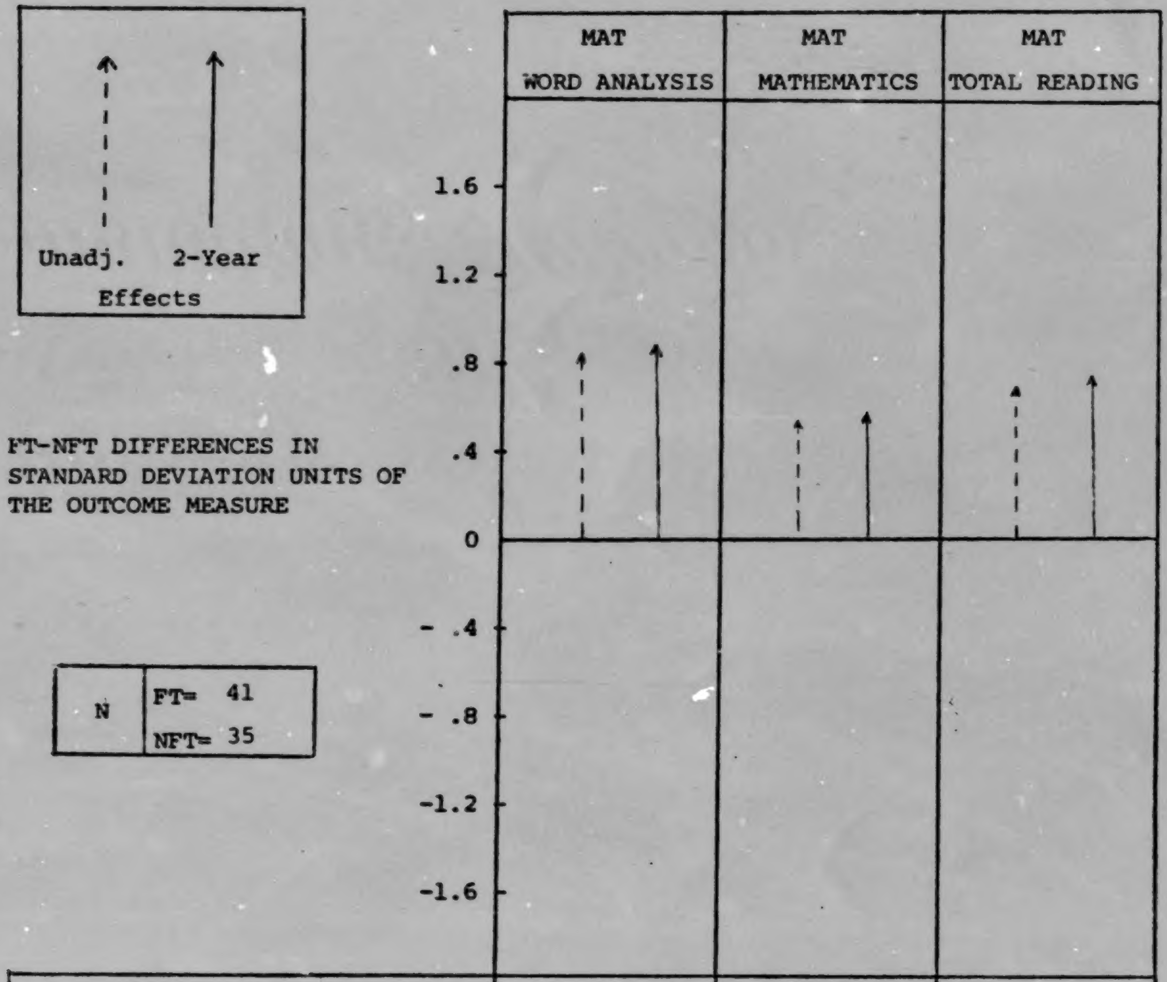
FIGURE VI-3 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FAR WEST SITE NC1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.128	- .540	8.474
	2-Year	.095	- 1.756	6.542
S.E. = Standard Error of B	Unadjusted	.041	1.831	2.971
	2-Year	.030	1.368	2.120
F = Significance Statistic	Unadjusted	9.747*	.087	8.134*
	2-Year	10.030*	1.649	9.523*
SD=Raw Score Standard Deviation		.234	9.994	16.745
B/SD = Effect in SD Units (See Graph)	Unadjusted	.547	- .054	.506
	2-Year	.406	- .176	.391

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-4 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FAR WEST SITE NC2



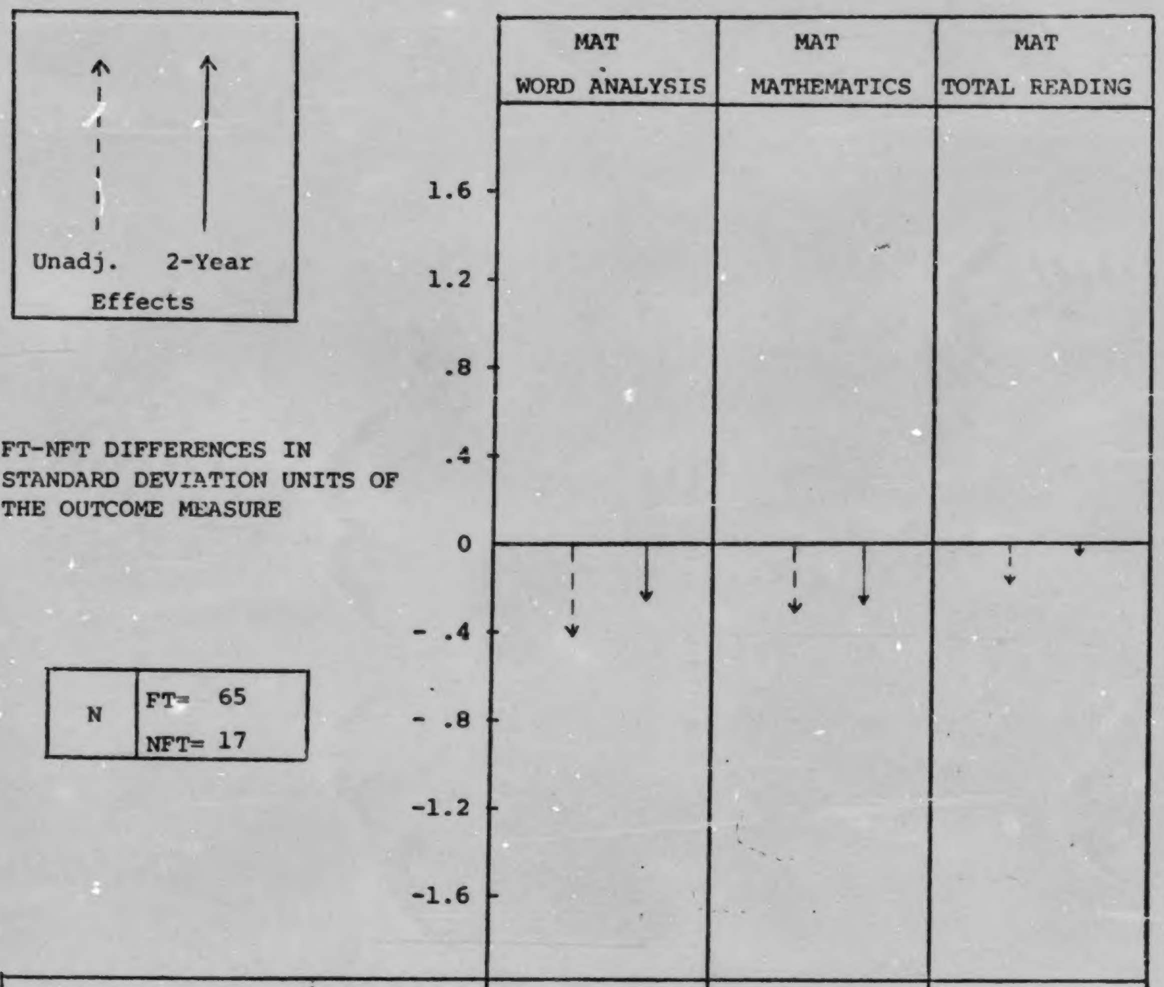
		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.178	5.842	10.017
	2-Year	.188	6.331	10.675
S.E. = Standard Error of B	Unadjusted	.044	2.399	3.093
	2-Year	.040	2.228	2.783
F = Significance Statistic	Unadjusted	16.366*	5.930*	10.491*
	2-Year	22.090*	8.077*	14.715*
SD=Raw Score Standard Deviation		.209	10.761	14.266
B/SD = Effect in SD Units (See Graph)	Unadjusted	.852	.543	.702
	2-Year	.900	.588	.748

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

In Far West's other three sites, the FT group is disadvantaged relative to the NFT group. In site NE (Figure VI-5), where the FT group shows a relatively small disadvantage on SES and a somewhat higher percentage of minority children, the FT/NFT differences at the end of first grade are small but favor NFT on Word Analysis and Mathematics. Site W1 (Figure VI-6) shows substantial differences favoring NFT in two year gains on all outcomes. In site W2 (Figure VI-7) the FT children are also disadvantaged relative to NFT, and the end of first grade effects show a difference favoring NFT only on Mathematics. As noted, this is a puzzling site where the indicators portray an inconsistent picture of the FT/NFT entry difference. However, the difference on level of father's education suggests that the FT group is more disadvantaged.

These site effects indicate that in at least some locales the Far West Program is producing substantially larger gains than those produced by the curriculum that a comparable group of NFT children experienced. In Far West's remaining sites, all involving a more or less severe mismatch of FT and NFT groups, the end of first grade effects are less favorable. However, the interpretation applied to the overall analysis may apply to the sites as well. In two of the three sites where the Far West FT group shows a disadvantage relative to NFT background and/or entry achievement variables, the FT group appears to be holding its own with respect to NFT. In these sites the FT group is either slightly behind NFT at the end of the first grade or the groups are performing at equal levels. In the remaining sites, the FT group is substantially behind NFT at the end of first grade and does not appear to be keeping up. Whether this site is truly an exception to the "keeping up" hypothesis and whether this hypothesis is at all valid cannot be determined with the available data. However, it seems preferable to be on the side of liberal interpretation of program effectiveness at this early stage because of the fallibility of method and the relative incompleteness of both the program and the data.

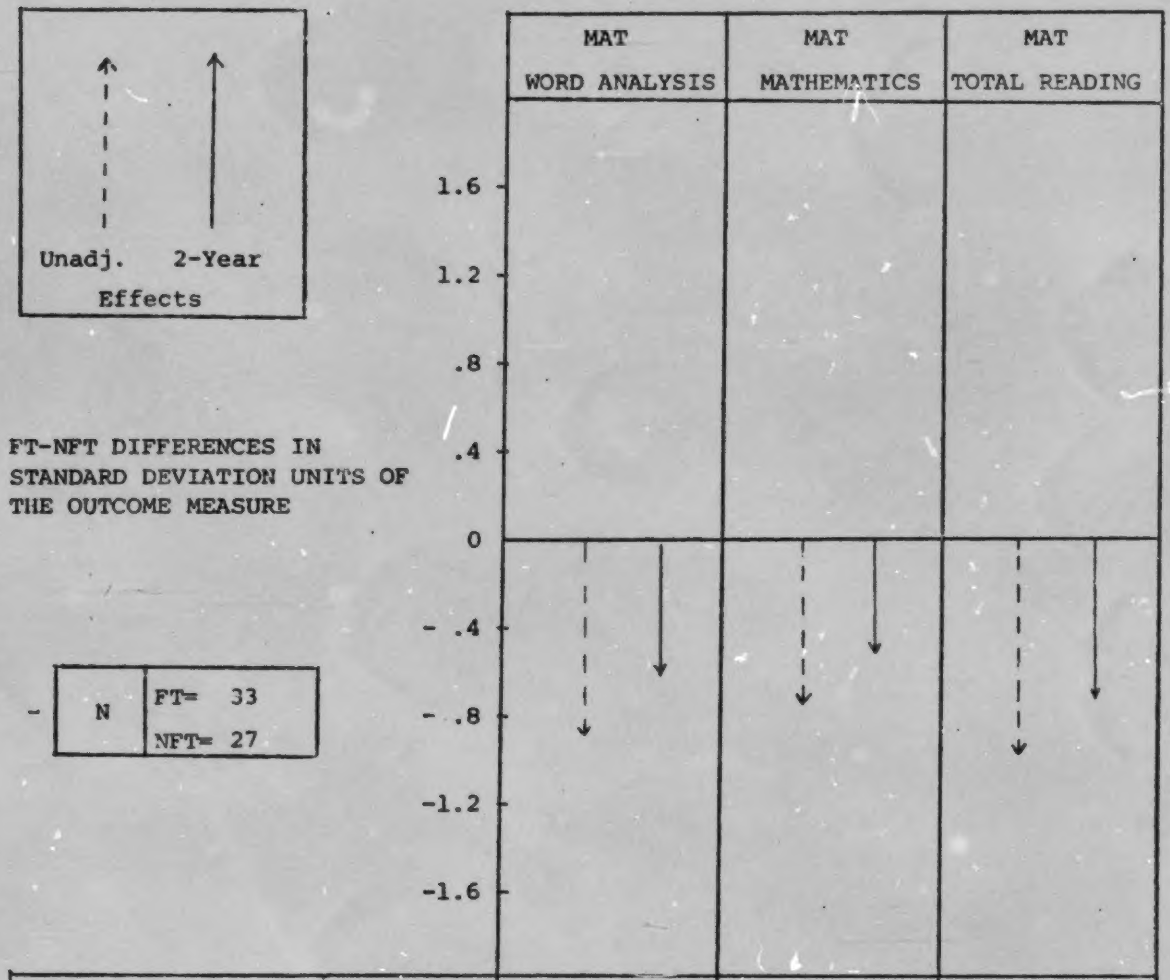
FIGURE VI-5 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FAR WEST SITE NE



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .086	- 3.009	- 2.151
	2-Year	- .052	- 2.524	- .432
S.E. = Standard Error of B	Unadjusted	.055	2.580	3.060
	2-Year	.050	2.450	2.715
F = Significance Statistic	Unadjusted	2.446	1.360	.494
	2-Year	1.082	1.061	.025
SD=Raw Score Standard Deviation		.203	9.491	11.197
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .424	- .317	- .192
	2-Year	- .256	- .266	- .039

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

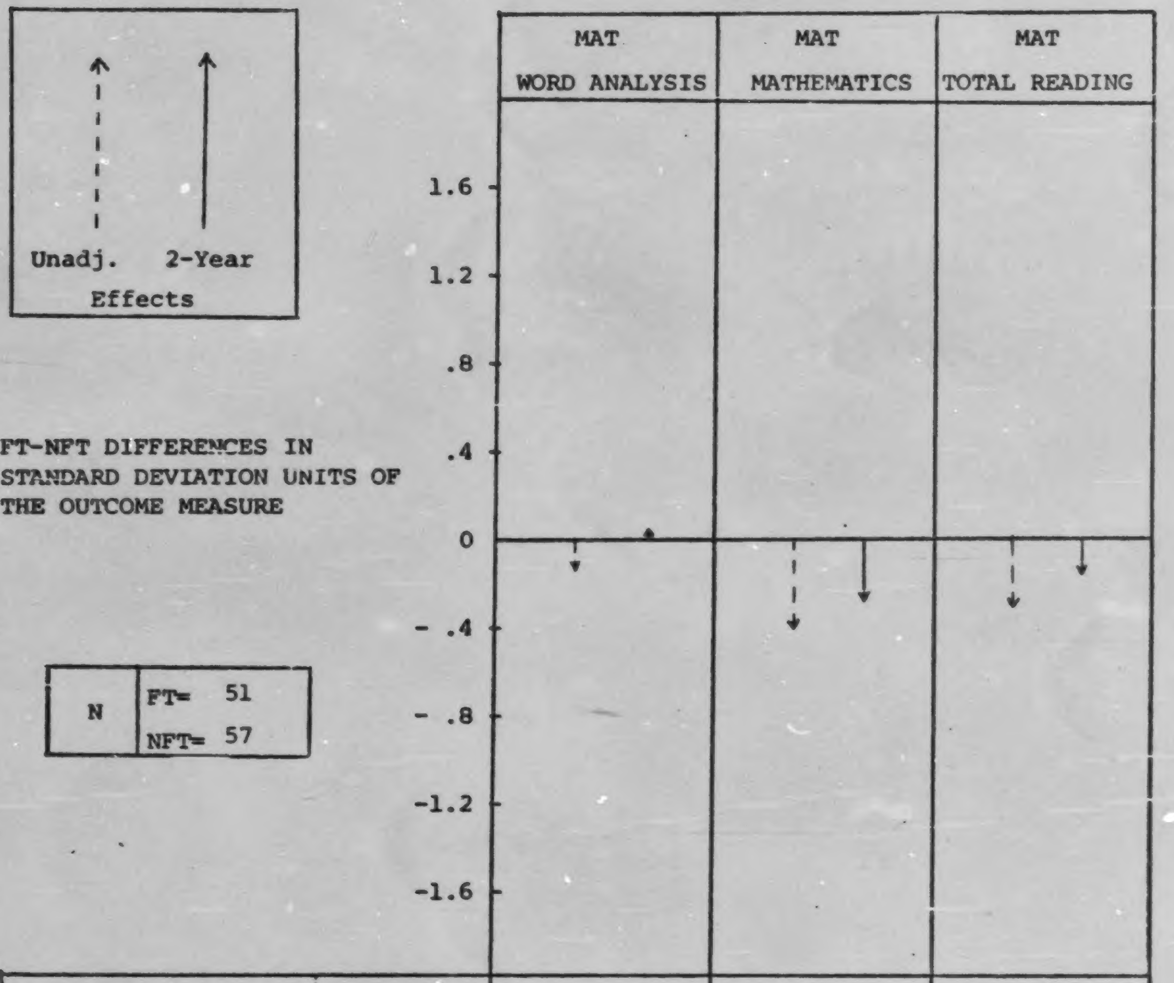
FIGURE VI-6 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FAR WEST SITE W1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .212	- 8.118	-11.640
	2-Year	- .146	- 5.559	- 8.662
S.E. = Standard Error of B	Unadjusted	.054	2.532	2.674
	2-Year	.042	2.037	2.162
F = Significance Statistic	Unadjusted	15.413*	10.278*	18.949*
	2-Year	12.083*	7.447*	16.048*
SD=Raw Score Standard Deviation		.233	10.495	11.768
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .910	- .774	- .989
	2-Year	- .627	- .530	- .736

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

FIGURE VI-7 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FAR WEST SITE W2



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .033	- 5.168	- 5.622
	2-Year	.008	- 3.542	- 2.766
S.E. = Standard Error of B	Unadjusted	.045	2.329	3.346
	2-Year	.036	1.988	2.657
F = Significance Statistic	Unadjusted	.537	4.924*	2.822
	2-Year	.049	3.176	1.084
SD=Raw Score Standard Deviation		.232	12.304	17.506
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .142	- .420	- .321
	2-Year	.034	- .288	- .158

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

6.2 UNIVERSITY OF ARIZONA, TUSCON EARLY EDUCATIONAL MODEL

The University of Arizona program focuses upon four general areas of development: language, reasoning abilities, motivation, and social arts and skills. The classroom environment is designed to reflect the child's home and community environment, so that skills and concepts may be learned in a natural, functional setting. One-to-one adult/child interactions and small group activities are utilized to individualize instruction and to help develop effective social interaction and communication skills.

6.2.1 Overall 2-Year Effects

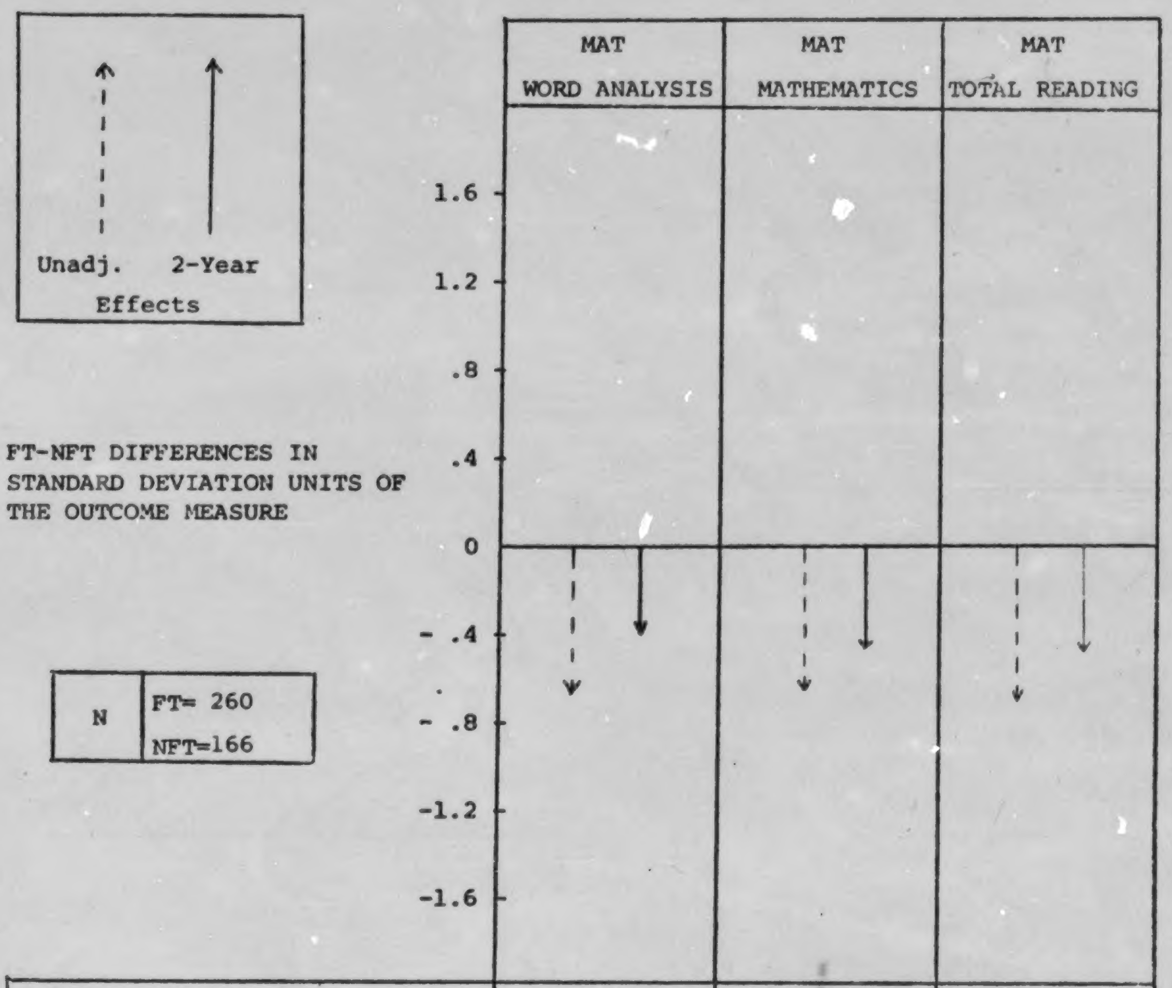
The FT/NFT differences for Arizona, shown in Figure VI-8 , indicate that the FT group is substantially behind NFT on all outcome measures. The Arizona FT and NFT groups are mismatched on SES and entry scores. This is indicated by the substantial differences between the adjusted and unadjusted scores. The differences at the end of first grade suggest that the FT children are falling increasingly further behind as grade school progresses. This trend is supported by the fact that the FT group was approximately .35 standard deviations behind NFT on the WRAT at entry into kindergarten and is somewhat further behind on the first grade measures.

6.2.2 Site Descriptions

Five sites met the criteria for inclusion in the evaluation of the University of Arizona model. These sites are scattered geographically, and their students differ in a number of ways. Four of the five sites have at least one characteristic in common; in these the FT groups are considerably poorer than their NFT counterparts.

The five Arizona sites differ greatly in size and location. Two, located in the Northeast, are a small town (NE1) and a large city (NE2). Two, located in the North Central region are a medium-sized city (NC1) and a large city (NC2). The remaining site is a small Southern town (S). The pupil and family characteristics for the sites are shown in Tables VI-6 through VI-10.

FIGURE VI-8 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
ARIZONA



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .178	- 8.698	-12.682
	2-Year	- .114	- 5.694	- 8.380
S.E. = Standard Error of B	Unadjusted	.026	1.195	1.691
	2-Year	.019	.895	1.254
F = Significance Statistic	Unadjusted	48.530*	52.976*	56.259*
	2-Year	35.383*	40.456*	44.694*
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .668	- .664	- .723
	2-Year	- .427	- .435	- .478

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

In the small Southern community (S, Table VI-6), the FT/NFT groups have different ethnic compositions (FT = 40% black, 60% white; NFT = 98% white). On measures of parental education, family income, and father's occupation, the NFT group exceeds the FT group by a substantial margin. In addition, the NFT children in this community are far more apt to come from intact families (FT = 54%; NFT = 94%).

In the medium-sized North Central city (NC1, Table VI-7), both FT and NFT children are predominantly White; in the large North Central city (NC2, Table VI-8), both groups contain roughly the same proportion of Black and White children, mostly White. However, in both communities, the FT children are much poorer than the NFT children. Furthermore, the average FT parent did not graduate from high school whereas the average NFT parent did. In NC1, both FT/NFT groups come from predominantly intact families while in NC2, the FT group has a smaller percentage of intact families (FT = 65%; NFT = 83%).

In the two Northeastern communities both FT/NFT students are predominantly Black. In NE1 (Table VI-9), the FT group is well below the NFT group on parents' education, the FT group averaging less than an eighth grade education. The FT group lags behind NFT on family income and occupation as well. In both groups only about half of the children come from homes with both a mother and father present.

Arizona is thus dealing with sites where the FT children are in general substantially disadvantaged relative to their NFT comparison groups.

6.2.3 By Site 2-Year Effects

At the end of first grade, the Arizona FT children are behind their NFT comparison groups by substantial margins on all outcomes (See Figures VI-9 through VI-13). Although the size of the NFT advantage at the end of first grade is smallest in the site where the FT and NFT groups are well matched, site NE2, even here the FT group is behind on all measures.

Arizona appears to be producing a consistent pattern of losses at the end of first grade. The NFT children are moving far ahead of the children in the Arizona program. Regardless of the comparability of the FT and NFT group, this pattern is evident in both the overall analysis as well as in each of the sites. Although the Arizona children may blossom in the second or third grade, the results suggest that these children are falling progressively further behind their more advantaged peers following the traditional divergent pattern of achievement performance.

TABLE VI-6

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR ARIZONA SITE S

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.44	1.13	36	6.09	1.08	38
Father's Education	5.44	1.47	23	6.54	1.19	37
Income	8.58	6.86	36	17.17	6.31	36
Occupation	15.56	1.93	36	13.55	2.44	38
Family Structure						
Mother and Father Present	54%	-	20	94%	-	36
Mother, not Father, Present	35	-	13	3	-	1
Other	11	-	4	3	-	1
TOTAL			37			38
Entry Level						
WRAT	24.82	13.30	39	36.37	13.11	46
PPVT	43.75	11.72	39	54.37	9.61	46
Preschool Experience						
Attended Headstart or other Preschool	58%	-	28	15%	-	7
TOTAL			48			46
Ethnicity						
Black	40%	-	16	0%	-	0
White	60	-	23	98	-	45
Spanish	0	-	0	0	-	0
Other	0	-	0	2	-	1
TOTAL			39			46

TABLE VI-7

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR ARIZONA SITE NC1

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.80	.80	51	6.14	.94	43
Father's Education	5.70	1.38	40	6.16	1.30	38
Income	11.96	7.18	51	16.51	6.36	43
Occupation	15.46	1.96	50	14.63	2.46	43
Family Structure						
Mother and Father Present	78%	-	40	85%	-	37
Mother, not Father, Present	20	-	10	11	-	5
Other	2	-	1	4	-	2
TOTAL			51			44
Entry Level						
WRAT	33.50	13.55	140	36.49	13.93	92
PPVT	2.84	8.24	140	55.06	9.34	92
Preschool Experience						
Attended Headstart or other Preschool	22%	-	32	9%	-	9
TOTAL			146			96
Ethnicity						
Black	8%	-	11	4%	-	4
White	86	-	121	93	-	86
Spanish	3	-	4	3	-	2
Other	3	-	4	0	-	0
TOTAL			140			92

TABLE VI-8

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR ARIZONA SITE NC2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.68	1.02	54	6.10	1.14	57
Father's Education	5.92	1.30	38	6.71	1.34	49
Income	11.16	6.97	56	15.17	7.62	54
Occupation	15.09	2.06	54	13.25	2.87	56
Family Structure						
Mother and Father Present	65%	-	37	83%	-	48
Mother, not Father, Present	25	-	14	13	-	8
Other	10	-	6	4	-	2
TOTAL			57			58
Entry Level						
WRAT	29.93	12.05	58	40.05	14.90	57
PPVT	47.59	10.14	58	53.74	10.31	57
Preschool Experience						
Attended Headstart or other Preschool	52%	-	33	56%	-	10
TOTAL			63			18
Ethnicity						
Black	32%	-	19	23%	-	13
White	56	-	32	67	-	38
Spanish	12	-	7	5	-	3
Other	0	-	0	5	-	3
TOTAL			58			57

TABLE VI-9

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR
ARIZONA SITE NE1

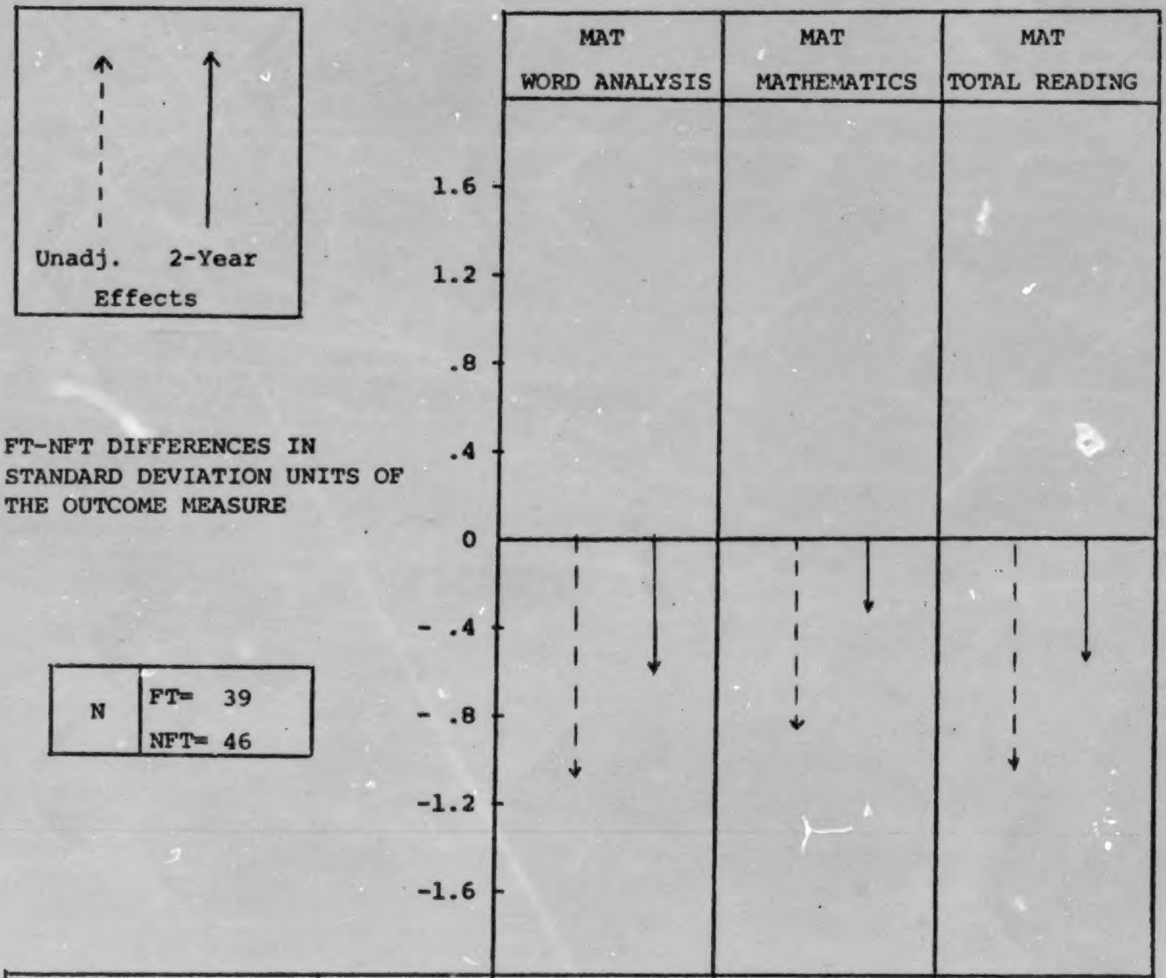
	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	4.88	1.49	68	5.68	.91	34
Father's Education	4.58	1.38	31	5.32	1.20	19
Income	9.19	5.60	73	11.95	7.10	37
Occupation	16.03	1.47	60	15.47	2.02	36
Family Structure						
Mother and Father Present	44%	-	32	45%	-	17
Mother, not Father, Present	48	-	35	41	-	16
Other	8	-	6	14	-	5
TOTAL			73			38
Entry Level						
WRAT	26.42	11.20	64	27.93	12.45	27
PPVT	38.45	15.38	64	38.04	14.47	27
Preschool Experience						
Attended Headstart or other Preschool	97%	-	73	missing		
TOTAL			75	missing		
Ethnicity						
Black	65%	-	42	66%	-	18
White	10	-	6	26	-	7
Spanish	25	-	16	8	-	2
Other	0	-	0	0	-	0
TOTAL			64			27

TABLE VI-10

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR ARIZONA SITE NE2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.32	1.04	34	5.29	.77	17
Father's Education	5.15	1.05	34	missing		
Income	9.12	5.93	60	10.17	6.52	18
Occupation	15.71	1.49	59	15.41	2.00	17
Family Structure						
Mother and Father Present	60%	-	37	40%	-	8
Mother, not Father, Present	35	-	22	50	-	10
Other	5	-	3	10	-	2
TOTAL			62			20
Entry Level						
WRAT	29.85	13.45	66	28.25	13.20	20
PPVT	34.71	13.62	66	37.50	13.34	20
Preschool Experience						
Attended Headstart or other Preschool	59%	-	28	missing		
TOTAL			47			
Ethnicity						
Black	97%	-	64	100%	-	20
White	0	-	0	0	-	0
Spanish	3	-	2	0	-	0
Other	0	-	0	0	-	0
TOTAL			66			20

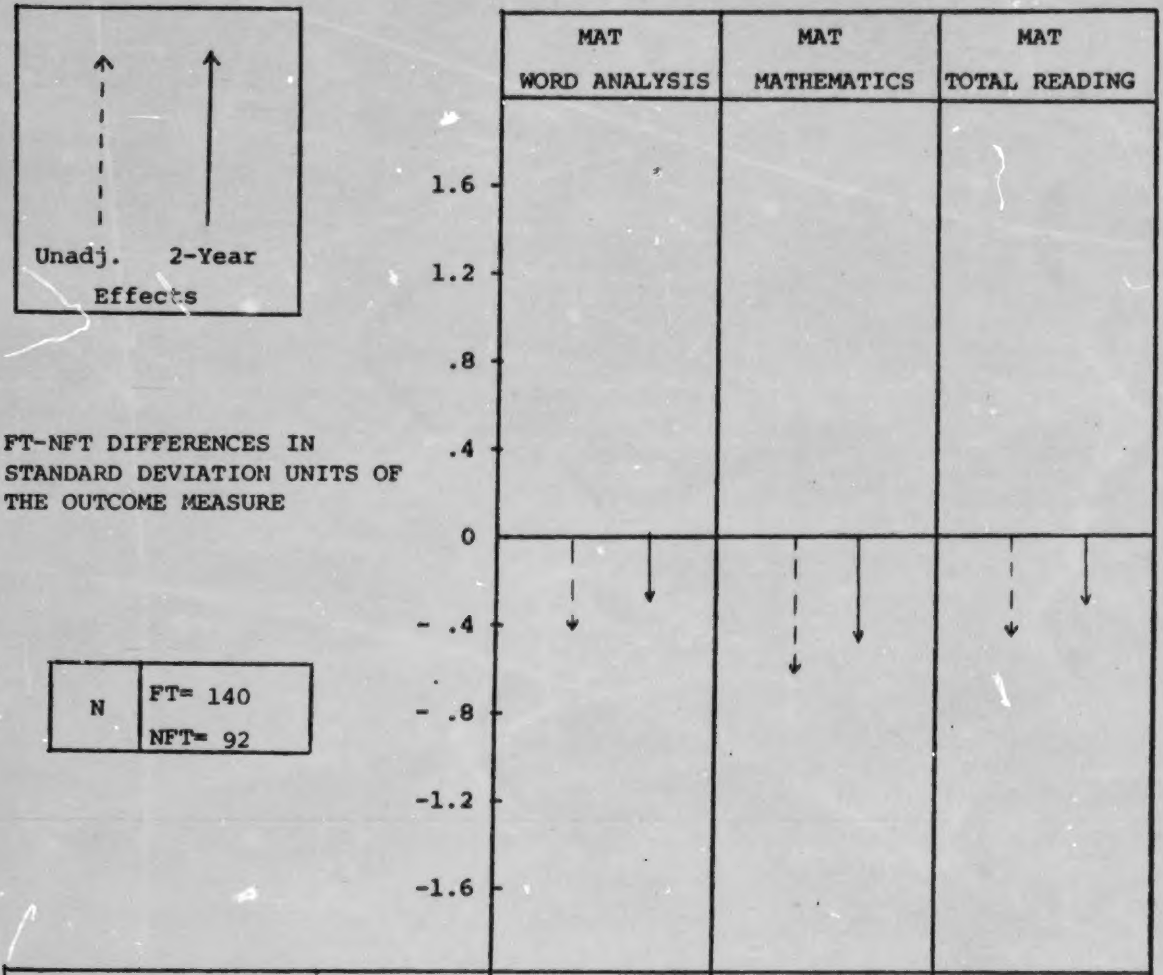
FIGURE VI-9 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
ARIZONA SITE S



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .292	-11.868	-19.333
	2-Year	- .167	- 4.431	-10.115
S.E. = Standard Error of B	Unadjusted	.048	2.678	3.314
	2-Year	.041	2.113	2.638
F = Significance Statistic	Unadjusted	37.003*	19.643*	34.036*
	2-Year	16.589*	4.397*	14.700*
SD=Raw Score Standard Deviation		.265	13.601	17.970
B/SD = Effect in SD Units (See Graph)	Unadjusted	- 1.102	- .873	- 1.076
	2-Year	- .630	- .326	- .563

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

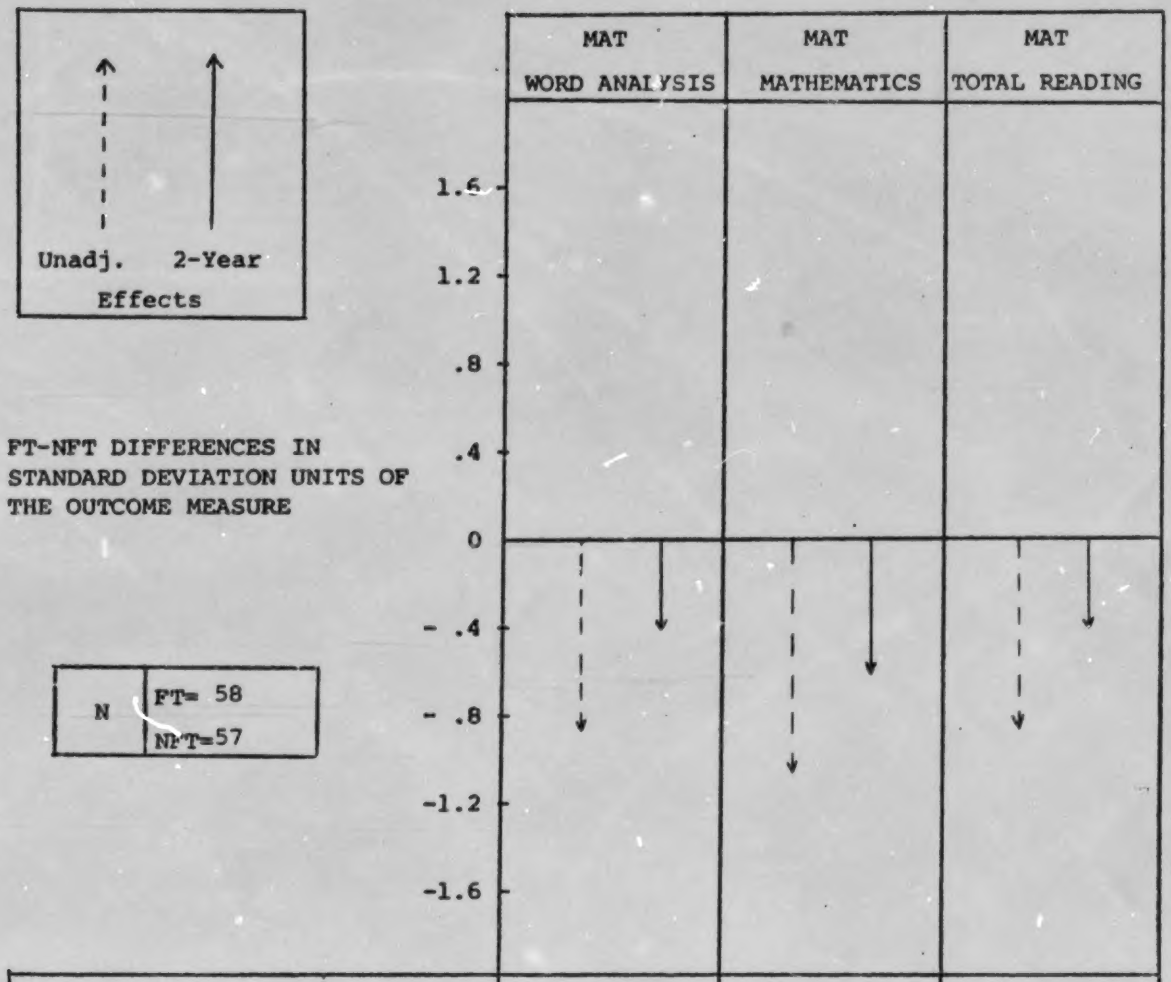
FIGURE VI-10 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
ARIZONA SITE NC1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .101	- 7.906	- 8.604
	2-Year	- .070	- 6.133	- 5.802
S.E. = Standard Error of B	Unadjusted	.031	1.593	2.436
	2-Year	.024	1.162	1.746
F = Significance Statistic	Unadjusted	10.615*	24.631*	12.475*
	2-Year	8.509*	27.857*	11.042*
SD=Raw Score Standard Deviation		.236	12.462	18.597
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .428	- .634	- .463
	2-Year	- .297	- .492	- .312

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

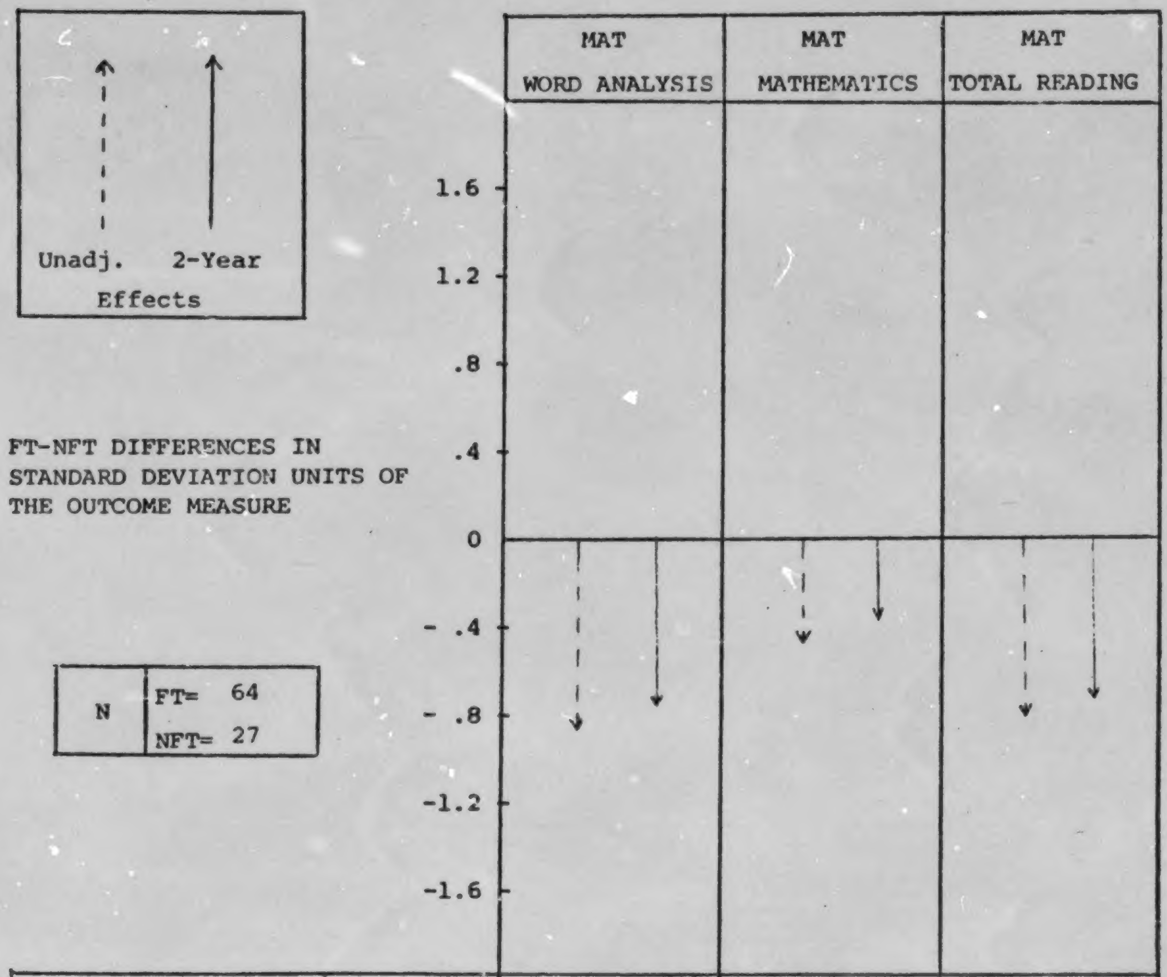
FIGURE VI-11 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
ARIZONA SITE NC2



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .224	-14.719	-15.268
	2-Year	- .107	- 8.647	- 7.270
S.E. = Standard Error of B	Unadjusted	.044	2.168	2.908
	2-Year	.036	1.861	2.399
F = Significance Statistic	Unadjusted	25.918*	46.090*	27.562*
	2-Year	8.833*	21.585*	9.181*
SD=Raw Score Standard Deviation		.258	13.733	17.313
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .868	- 1.072	- .882
	2-Year	- .415	- .630	- .420

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

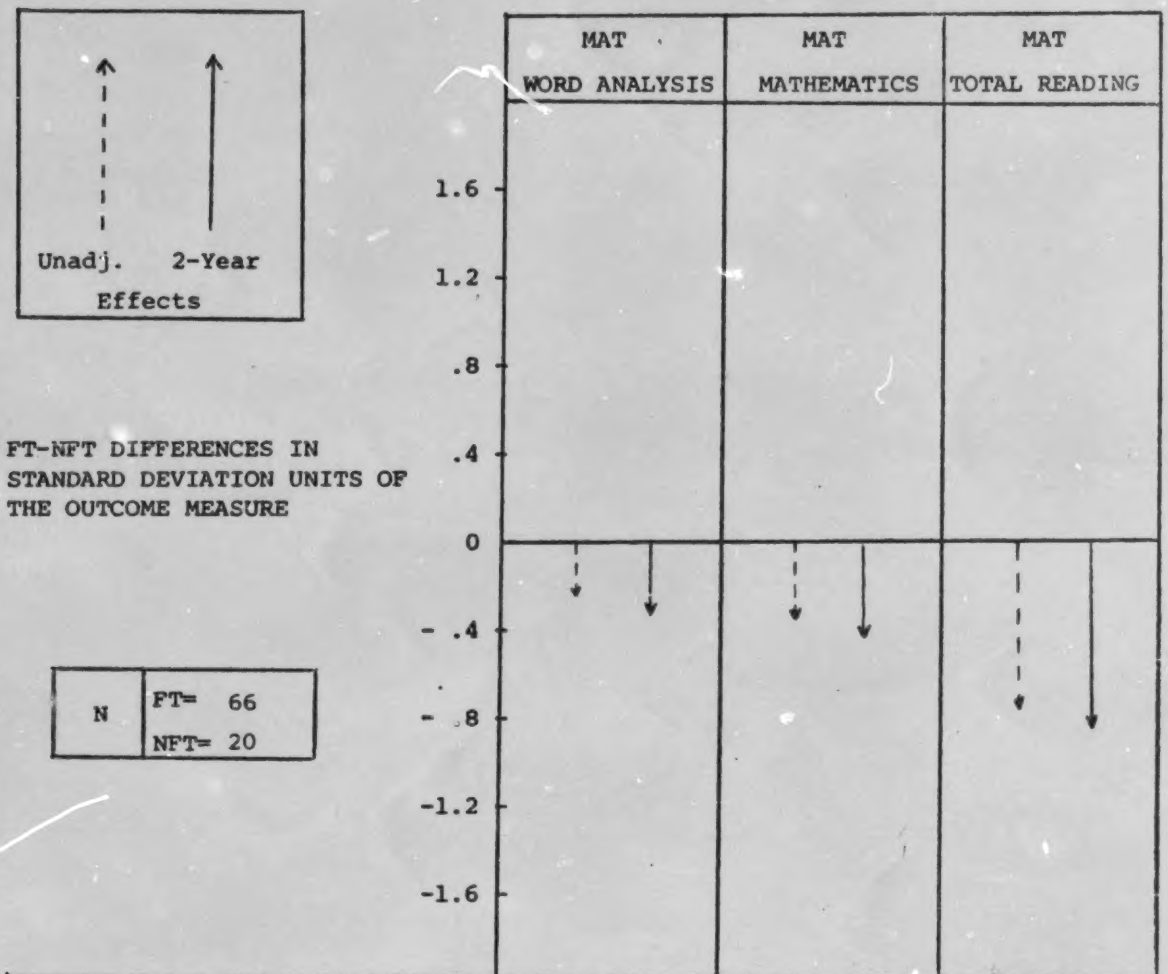
FIGURE VI-12 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
ARIZONA SITE NE1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .232	- 5.271	-12.922
	2-Year	- .208	- 4.310	-11.634
S.E. = Standard Error of B	Unadjusted	.056	2.544	3.321
	2-Year	.045	2.071	2.660
F = Significance Statistic	Unadjusted	17.164*	4.293*	15.140*
	2-Year	21.363*	4.331*	19.132*
SD=Raw Score Standard Deviation		.266	11.285	15.568
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .872	- .467	- .830
	2-Year	- .782	- .382	- .747

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

FIGURE VI-13 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
ARIZONA SITE NE2



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .062	- 3.196	-11.647
	2-Year	- .081	- 3.860	-12.705
S.E. = Standard Error of B	Unadjusted	.063	2.257	3.593
	2-Year	.049	1.802	2.871
F = Significance Statistic	Unadjusted	.968	2.005	10.511*
	2-Year	2.732	4.588*	19.581*
SD=Raw Score Standard Deviation		.247	8.894	14.843
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .251	- .359	- .785
	2-Year	- .328	- .434	- .856

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's educatio
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

6.3 BANK STREET COLLEGE APPROACH

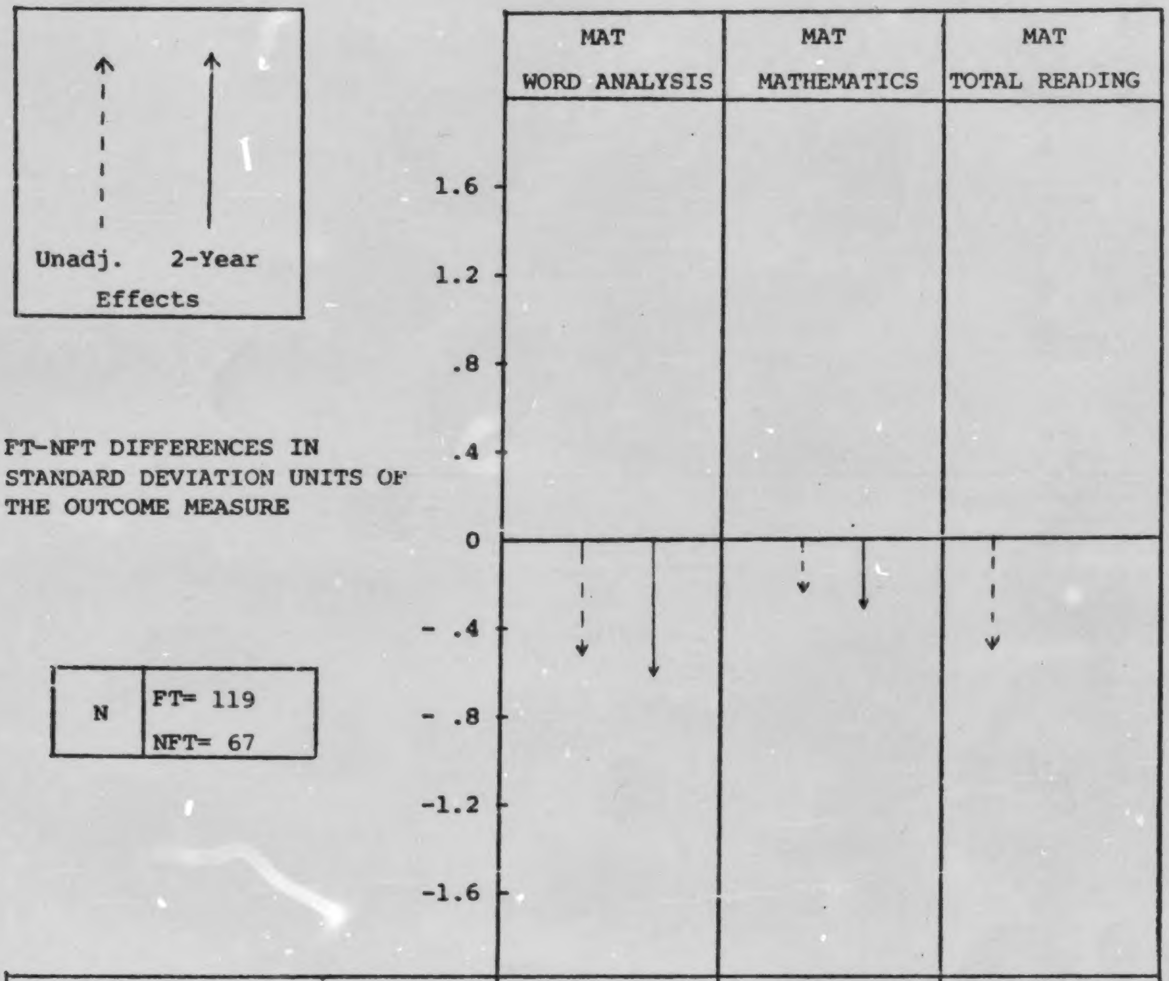
The Bank Street approach is designed to change the school system to meet the developmental needs of children. Heavy emphasis is placed on teacher training to help teachers organize the classroom environment for self-directed learning and to plan events to meet the needs of children. The individualized, flexible curriculum is designed not only to help children acquire basic skills, but also to help children master how to learn. Creativity and self-expression are also important program goals.

An evaluation of the Bank Street program is severely handicapped by drastic losses of sites and children. Of the four sites available for study in the first grade year, two, both big cities, lost large numbers of children because of family mobility. A further reduction in sample size resulted from incomplete test and parent interview data. The numbers of children for whom complete data are available at these two sites are not sufficient to permit any kind of analysis, although other data allow us to draw some inferences about the effectiveness of the program in these settings. The two remaining sites with large enough numbers to permit site level analysis are both New England towns whose demographic and educational characteristics are very unlike the other Bank Street FT sites.

6.3.1 Overall 2-Year Effects

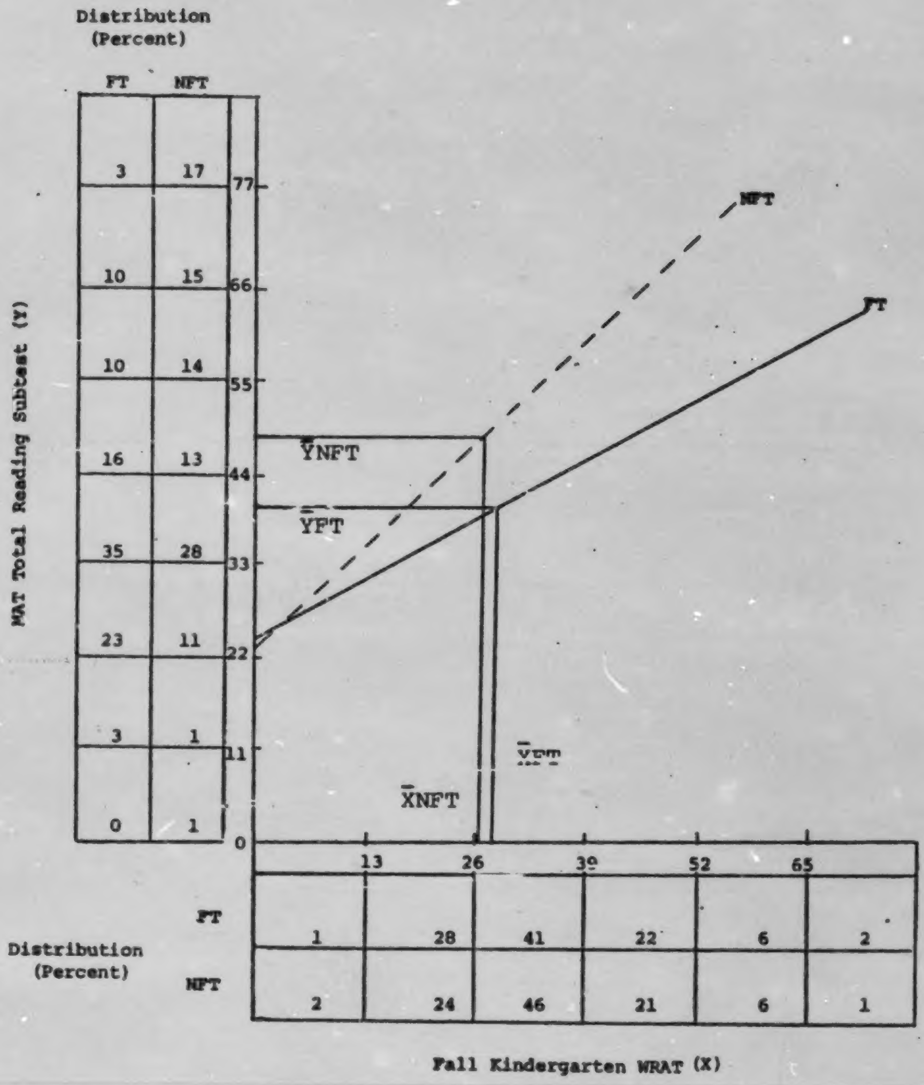
The overall differences between Bank Street's FT and NFT groups, shown in Figure VI-14, indicate that the Bank Street children are substantially behind their NFT comparison group on Word Analysis and only slightly behind on Mathematics. For Total Reading, only the unadjusted difference is shown since the relationships between reading posttest scores and entry WRAT scores in FT and NFT are substantially different. These heterogeneous relationships, shown in Figure VI-15, suggest that higher achieving NFT children gained somewhat more than higher achieving FT children. The figure also shows that the FT and NFT groups are about equal on entry WRAT and that the NFT group is substantially ahead, .51 standard deviations, on the Total Reading test at the end of first grade. The substantial disadvantage of the FT children in Total Reading as well as Word Analysis suggests that the Bank Street model is not particularly effective in these achievement areas with these children.

FIGURE VI-14 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
BANK STREET



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .140	- 3.121	- 8.947
	2-Year	- .166	- 4.164	
S.E. = Standard Error of B	Unadjusted	.038	1.760	2.479
	2-Year	.028	1.435	
F = Significance Statistic	Unadjusted	13.878*	3.145	13.026*
	2-Year	34.516*	8.420*	
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .526	- .238	- .510
	2-Year	- .622	- .318	

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq .05$



	FT	NFT
X	28.32	26.52
Y	39.28	48.22
b	.83	1.17
SE (b)	.09	.14
n	119	67

Figure VI-15

Heterogeneous Regressions:

Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT

Sponsor: Bank Street

Outcome: MAT Total Reading Subtest

6.3.2 Site Descriptions

As noted, only two of Bank Street's sites are represented in the analysis. The demographic and pupil characteristics of the children in these sites are shown in Tables VI-11 and VI-12. These two sites are both New England communities; the larger site (NE 2) has a population of 100,000 persons; and the other site (NE 1) less than 15,000. In both, the FT/NFT children are virtually all white.

In the smaller of the two communities (NE1, Table VI-11), the FT group has a much lower family income, has by far fewer intact families (FT = 52%; NFT = 90%), and is considerably behind the NFT group on kindergarten entry level scores. In the larger New England city (NE 2, Table VI-12), both FT and NFT children come from low SES families. The two groups are roughly equal in terms of family income and parental education, both groups averaging less than eight years. However, the FT group has a smaller percentage of intact families (FT = 57%; NFT = 81%).

While these two communities differ in some ways, they have something important in common. In both, many of the lower SES children are enrolled in the FT program.

6.3.3 By Site 2-Year Effects

In both of Bank Street's sites the FT children are substantially behind the NFT comparison children on Word Analysis and Total Reading. In site NE1, the site where the FT children show a substantial disadvantage on entry tests, they are behind on Mathematics as well (See Figures VI-16 and VI-17).

The Bank Street Program does not produce substantial gains over the kindergarten and first grade years in these sites, nor does the program appear to be particularly effective overall.

We are left with only a partial answer to the question of Bank Street's efficacy at the end of first grade. It remains to be seen whether the Sponsor produces such effects in different contexts, although there is some evidence to suggest this is so.

TABLE VI-11

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR BANK STREET SITE NE1

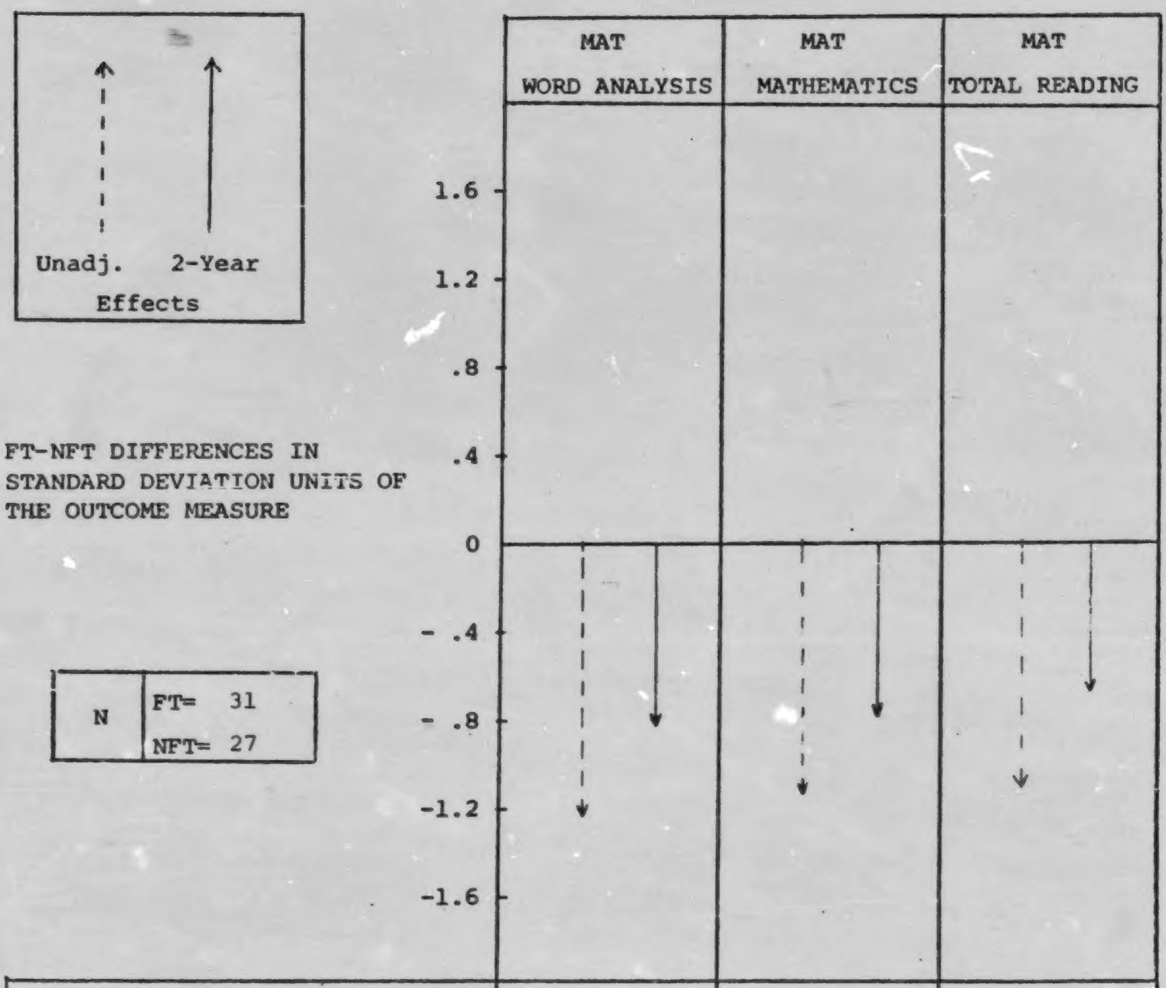
	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.83	.90	29	6.17	.89	29
Father's Education	5.68	1.53	19	5.93	1.04	27
Income	10.94	6.81	33	16.72	5.62	29
Occupation	15.48	2.02	33	14.34	2.51	29
Family Structure						
Mother and Father Present	52%	-	17	90%	-	26
Mother, not Father, Present	33	-	11	3	-	1
Other	15	-	5	7	-	2
TOTAL			33			29
Entry Level						
WRAT	28.64	12.78	31	37.15	14.96	27
PPVT	49.90	9.07	31	54.96	12.25	27
Preschool Experience						
Attended Headstart or other Preschool	missing			missing		
TOTAL	missing			missing		
Ethnicity						
Black	0%	-	0	0%	-	0
White	100	-	31	100	-	27
Spanish	0	-	0	0	-	-
Other	0	-	0	0	-	-
TOTAL			31			27

TABLE VI-12

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR BANK STREET SITE NE2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	4.69	1.07	55	3.67	1.45	45
Father's Education	4.56	1.33	34	3.60	1.72	37
Income	11.00	6.86	56	12.20	6.25	44
Occupation	14.38	1.48	55	16.30	1.81	46
Family Structure						
Mother and Father Present	57%	-	33	81%	-	37
Mother, not Father, Present	37	-	21	17	-	8
Other	6	-	3	2	-	1
TOTAL			57			46
Entry Level						
WRAT	24.34	10.66	71	24.43	10.11	51
PPVT	42.17	12.39	71	34.61	15.60	51
Preschool Experience						
Attended Headstart or other Preschool	51%	-	38	38%	-	13
TOTAL			74			34
Ethnicity						
Black	1%	-	1	0%	-	0
White	99	-	70	100	-	51
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			71			51

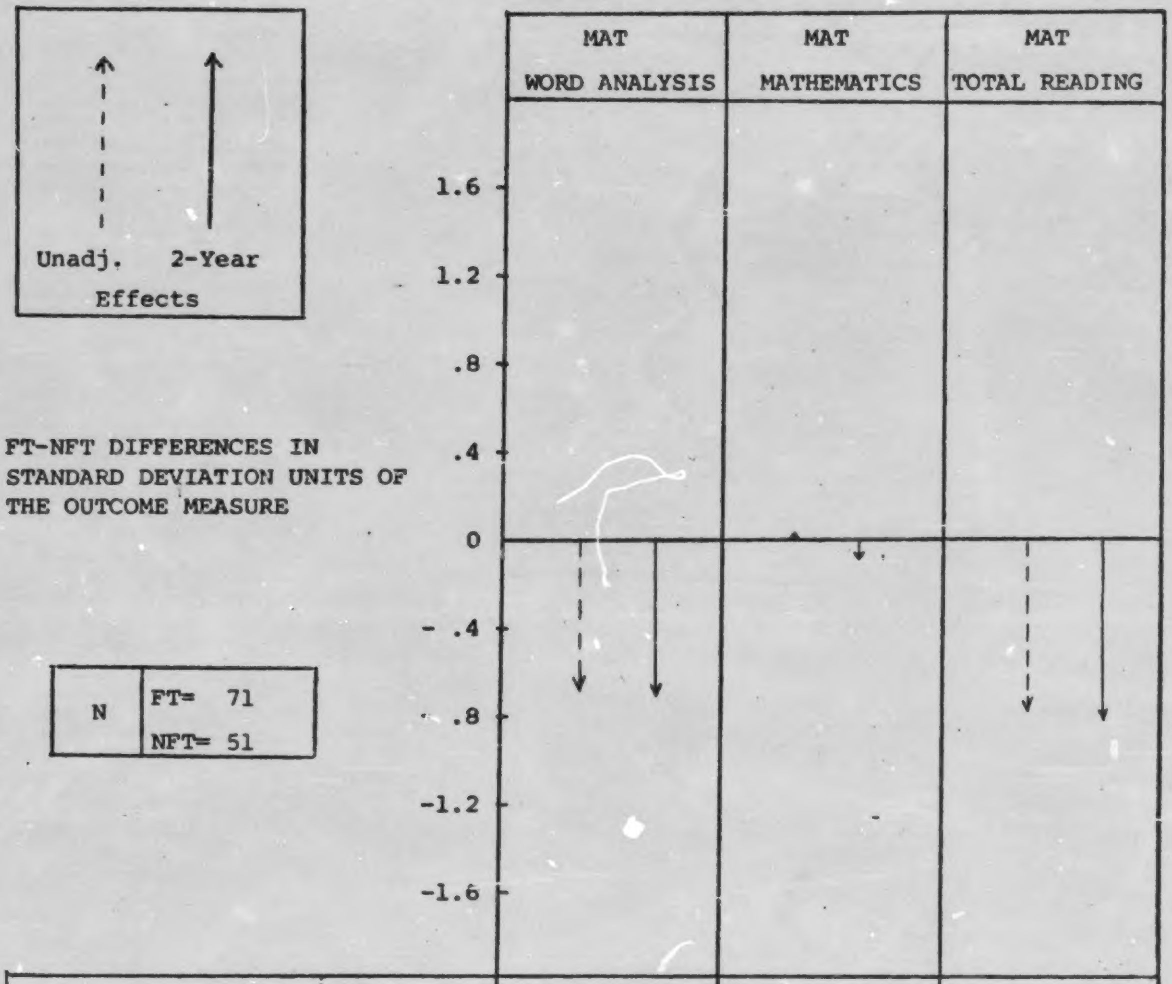
FIGURE VI-16 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
BANK STREET SITE NE1



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .375	-14.087	-22.239
	2-Year	- .255	- 9.727	-13.174
S.E. = Standard Error of B	Unadjusted	.062	2.648	4.310
	2-Year	.048	2.352	3.227
F = Significance Statistic	Unadjusted	36.578*	28.302*	26.626*
	2-Year	28.217*	17.106*	16.663*
SD=Raw Score Standard Deviation		.300	12.233	19.714
B/SD = Effect in SD Units (See Graph)	Unadjusted	- 1.250	- 1.152	- 1.128
	2-Year	- .850	- .795	- .668

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-17 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
BANK STREET SITE NE2



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .171	.030	-12.070
	2-Year	- .178	- .993	-12.617
S.E. = Standard Error of B	Unadjusted	.042	1.968	2.587
	2-Year	.036	1.662	2.173
F = Significance Statistic	Unadjusted	16.573*	.000	21.772*
	2-Year	24.443*	.376	33.710*
SD=Raw Score Standard Deviation		.245	10.678	15.257
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .698	.003	- .791
	2-Year	- .726	- .093	- .830

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

6.4 UNIVERSITY OF OREGON

The University of Oregon approach is to teach children the basic skills of reading, arithmetic, and language. It is based on the assumption that disadvantaged children can perform at "normal" levels of achievement when the instructional program builds on the skills children bring with them to school. The curriculum materials are programmed and sequenced. Teaching techniques are highly prescribed with heavy emphasis on structured small group instruction, quick paced question-and-answer periods, and the use of positive reinforcement to shape behavior.

6.4.1 Overall 2-Year Effects

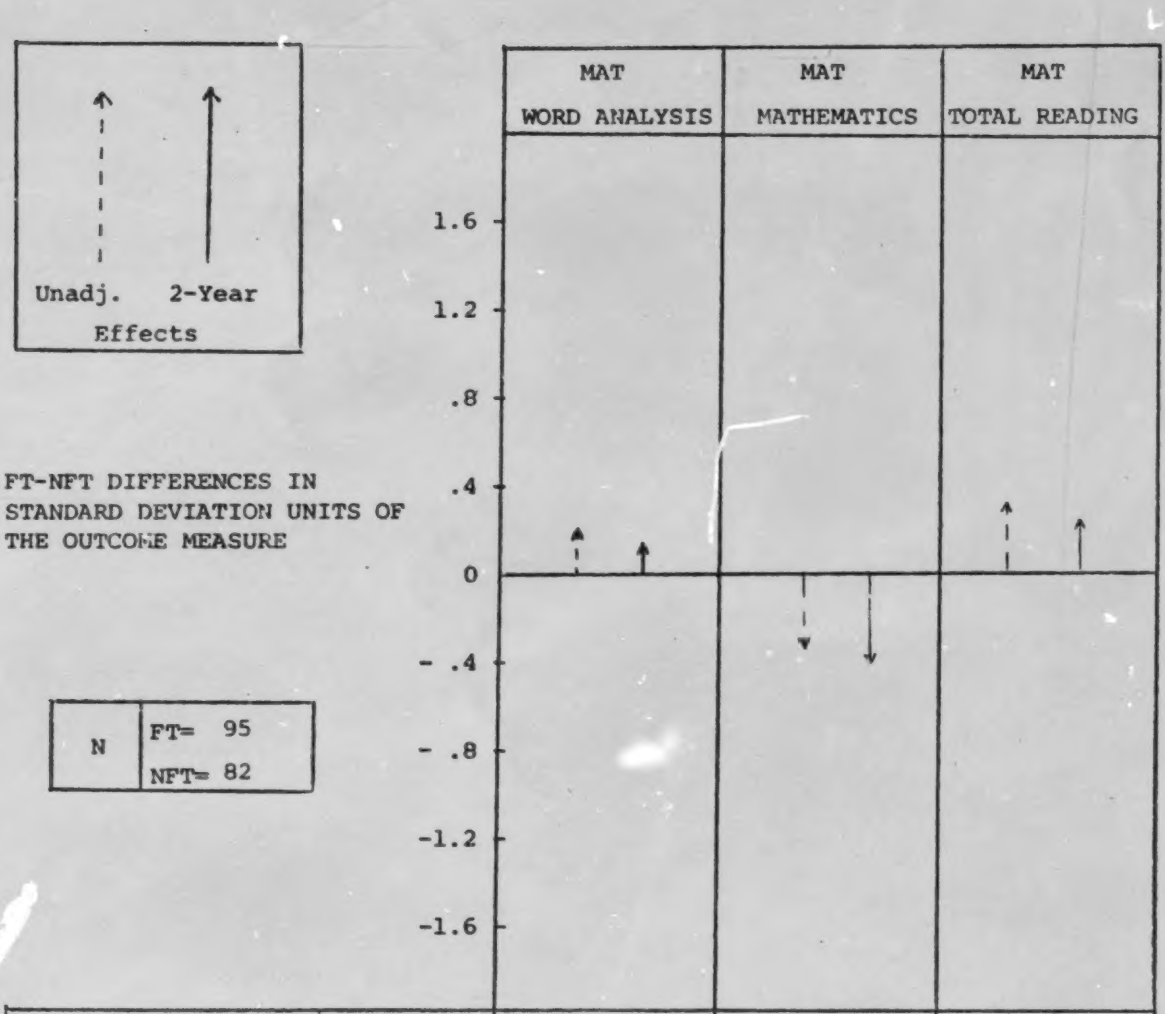
The FT/NFT differences for Oregon at the end of first grade, shown in Figure VI-18, indicate non-significant differences favoring FT between FT and NFT on Word Analysis and Total Reading but favoring NFT on Mathematics. Since the Oregon FT children comprise a lower income group with a high percentage of minority children and with a high percentage of children who attended preschool relative to their NFT comparisons, these results suggest that the FT children are keeping up with the more advantaged NFT group, at least in respect to Word Analysis and Total Reading.

The FT-favoring entry differences for Oregon on the WRAT and PPVT, which may represent an effect of preschool, an early curriculum effect, or an interaction of both, are somewhat problematic. The FT entry advantage produced an adjustment on end of first grade outcomes which is inconsistent with some of the background characteristics of the FT children. In light of this inconsistency, the overall position of FT relative to NFT, i.e., no difference, should be interpreted tenuously since all of the FT/NFT differences are marginal and since a small change in the nature of the adjustment for initial differences could change the pattern to one more favorable to Oregon.

6.4.2 Site Descriptions

Three sites are included in this analysis of the University of Oregon model. They are all cities of approximately 200,000 people, some-

FIGURE VI-18 END OF FIRST GRADE FT/NFT CONTRASTS TRACED 2-YEAR SAMPLE OREGON



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.057	- 4.620	5.600
	2-Year	.042	- 5.434	4.224
S.E. = Standard Error of B	Unadjusted	.035	1.862	2.722
	2-Year	.030	1.627	2.246
F = Significance Statistic	Unadjusted	2.711	6.156*	4.233*
	2-Year	1.913	11.153*	3.539
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	.215	- .353	.319
	2-Year	.157	- .415	.241

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

what diverse geographically, although they do not include any of the Southern communities with which this Sponsor works. The family and pupil entry characteristics of Oregon's FT and NFT children, shown in Tables VI-13 through VI-15, differ substantially both within and across sites.

The first site (NC1, Table VI-13) is a city in the North Central region of the country in which both the FT and NFT children are predominantly Black (FT = 86%; NFT = 83%). The FT group has lower family incomes than the NFT group and is somewhat lower on the level of parental education. There are slightly more intact families in the NFT group. The kindergarten entry level scores present a sharp contrast to this pattern of relative disadvantage on the part of the FT children; in site NC1, FT exceeds NFT on the fall kindergarten WRAT (FT = 27.96; NFT = 23.73) and on the PPVT (FT = 42.77; NFT = 40.42). Since 93% of the FT children compared with only 63% of the NFT children attended preschool programs, one might suspect that early school experience accounts for some part of the initial kindergarten advantages of the FT group. But we know from Vol. IA that the relationship between preschool and the fall kindergarten WRAT is not nearly strong enough to account for the large difference. Perhaps the University of Oregon program produces early achievement effects in the first weeks of school that the fall test battery comes late enough to detect.

The second site (NC2, Table VI-14) is also a North Central city, and, again, both the FT and NFT groups are predominantly Black (FT = 98%; NFT = 87%). The FT children are somewhat disadvantaged compared with their NFT counterparts; the family income for this FT group is lower than for the NFT group. Moreover, only 29% of the FT children come from families with both mother and father present, whereas 54% of the NFT children come from intact families.

As in site NC1, however, the relative position of the FT/NFT groups is reversed on entry level scores. The FT group exceeds the NFT group on the fall WRAT (FT = 33.69; NFT = 27.57) and the fall PPVT (FT = 43.80; NFT = 41.32). Once again, preschool attendance may have contributed to this fall achievement advantage. Fully 94% of the FT group versus only 20% of the NFT group attended a preschool program of some kind.

TABLE VI-13

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR OREGON SITE NC1

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.23	1.15	47	5.56	1.11	36
Father's Education	4.95	1.28	20	5.62	1.02	21
Income	7.12	5.23	49	10.42	7.77	33
Occupation	16.00	1.99	47	15.43	2.09	35
Family Structure						
Mother and Father Present	41%	-	21	56%	-	20
Mother, not Father, Present	49	-	25	41	-	15
Other	10	-	5	3	-	1
TOTAL			51			36
Entry Level						
WRAT	27.96	11.41	47	23.73	9.93	33
PPVT	42.77	12.65	47	40.42	12.12	33
Preschool Experience						
Attended Headstart or other Preschool	93%	-	30	63%	-	7
TOTAL			32			11
Ethnicity						
Black	86%	-	40	83%	-	27
White	12	-	6	17	-	6
Spanish	2	-	1	0	-	0
Other	0	-	0	0	-	0
TOTAL			47			33

TABLE VI-14

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR OREGON SITE NC2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.50	.80	54	5.41	.98	49
Father's Education	5.12	1.26	16	5.50	.97	30
Income	7.11	5.92	54	13.84	6.52	51
Occupation	15.51	1.84	47	15.89	1.11	45
Family Structure						
Mother and Father Present	29%	-	16	54%	-	28
Mother, not Father, Present	69	-	38	40	-	21
Other	2	-	1	6	-	3
TOTAL			55			52
Entry Level						
WRAT	33.69	10.78	49	27.57	9.76	47
PPVT	43.80	12.27	49	41.32	10.95	47
Preschool Experience						
Attended Headstart or other Preschool	94%	-	16	20%	-	10
TOTAL			17			51
Ethnicity						
Black	98%	-	48	87%	-	41
White	2	-	1	11	-	5
Spanish	0	-	0	2	-	1
Other	0	-	0	0	-	0
TOTAL			49			47

Before turning to the Northeastern site, it is important to note that in both of these communities the NFT schools receive a larger Title I allocation per pupil than the FT schools. In the case of NC2, the number of man days of professional support services per NFT child is also substantially greater than that for FT children. Thus, it appears that the NFT groups in these sites are also somewhat disadvantaged since they are eligible to receive compensatory services.

A third site (NE, Table VI-15)is a city in the Northeast, in which 76% of the FT children and 58% of the NFT children are Black. At this site, unlike the others, the FT group is substantially lower on both SES and kindergarten entry level WRAT scores than the NFT group. The NFT group exceeds the FT group on parents' level of education. The NFT mothers average slightly more than a high school education while the NFT fathers average some college. The FT mothers average less than complete high school, and there were no data available for FT fathers. The NFT group also exceeds the FT group on income level and on occupational status (lower values indicate higher status). In addition, there are more intact homes in the NFT group. The NFT group also exceeds the FT group by a substantial margin on fall WRAT (FT = 29.97; NFT = 37.75), but not on the PPVT scores (FT = 41.53; NFT = 42.85). The fact that 100% of the FT children versus only 15% of the NFT children attended preschool does not appear to have had a substantial impact on the relative achievement level of these two groups at entry to kindergarten. If this experience has an effect in the other Oregon communities, unmeasured site-specific characteristics have mediated the effect here.

Finally, this site differs from the two others in the concentration of Title I funds. The FT schools in this site receive a substantial amount of Title I funds, whereas the NFT schools receive little or none. It appears that in this site the FT group represents a distinct subgroup in the community, separated by a variety of demographic characteristics that are reflected in pretest scores.

TABLE VI-15

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR OREGON SITE NE

	PT			NPT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.28	.82	36	6.13	1.14	23
Father's Education	missing			7.07	1.07	14
Income	6.38	5.08	37	11.96	8.58	25
Occupation	16.24	.96	34	13.14	2.82	22
Family Structure						
Mother and Father Present	30%	-	11	50%	-	13
Mother, not Father, Present	67	-	25	38	-	10
Other	3	-	1	12	-	3
TOTAL			37			26
Entry Level						
WRAT	29.97	9.94	36	37.75	14.72	20
PPVT	41.53	9.47	36	42.84	17.92	20
Preschool Experience						
Attended Headstart or other Preschool	100%	-	36	15%	-	3
TOTAL			36			19
Ethnicity						
Black	76%	-	27	58%	-	12
White	24	-	9	42	-	8
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			36			20

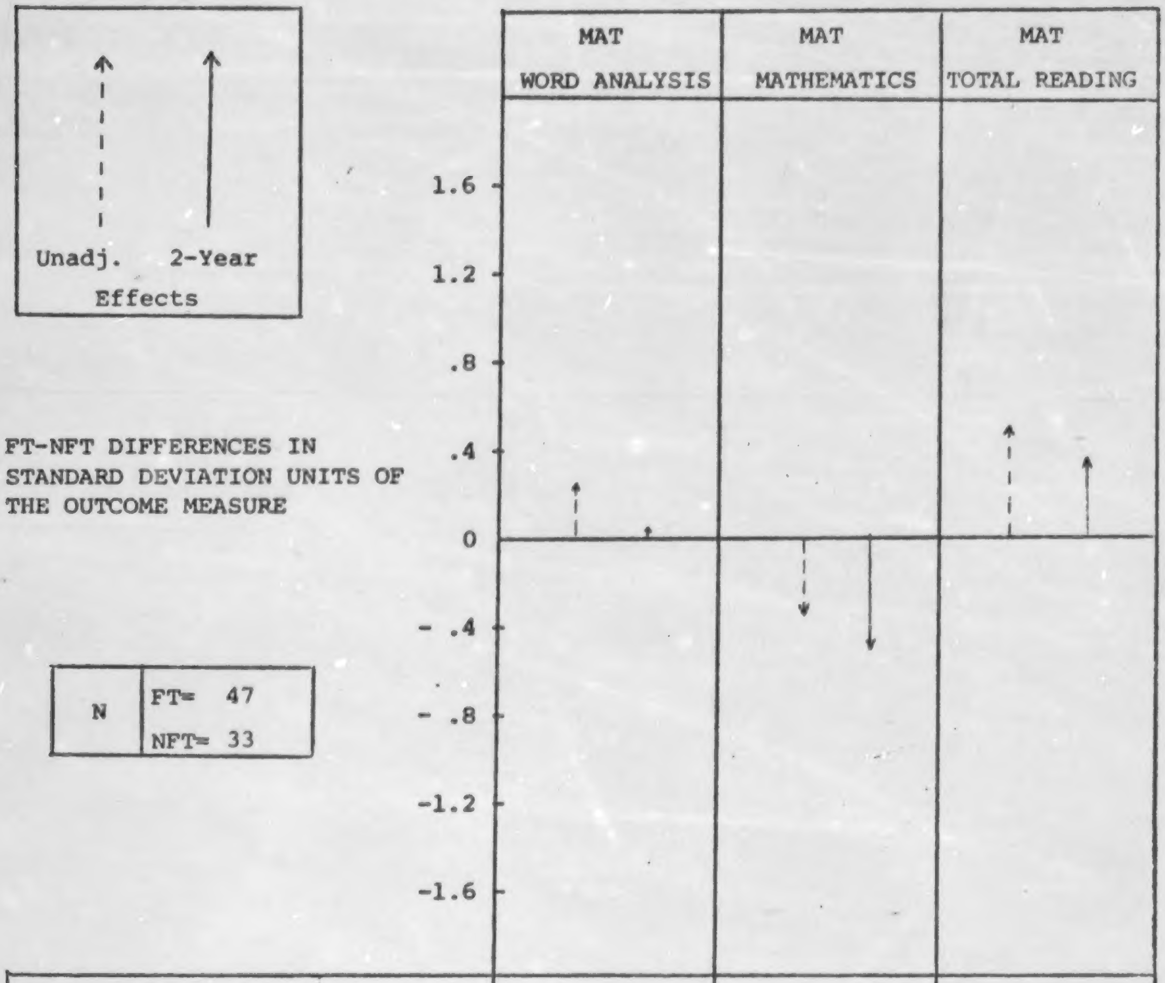
6.4.3 By Site 2-Year Effects

The FT/NFT differences for Oregon's sites, shown in Figures VI-19 through VI-21, are similar to the overall FT/NFT pattern. All three sites show FT-favoring differences on Total Reading, either no difference or an FT-favoring difference on Word Analysis, and either an NFT-favoring difference or no difference on Mathematics. Since the FT/NFT differences are relatively small in most cases, small changes in the adjustment for initial differences could change the pattern. However, since the FT children are in general somewhat disadvantaged relative to NFT on background factors, and since adjustment for initial differences for two of the sites is inconsistent with the FT disadvantage on background, any shift in effects would likely make the Oregon program look even more effective than the site profiles suggest.

In general, despite rather diverse sites and children, Oregon has produced consistent effects at the end of the first two years with Cohort III traced children. Total Reading and Word Analysis represent areas of equality between FT and NFT, whereas Mathematics is a domain in which the FT groups have fallen behind. This is a different picture of effects than those found at the end of kindergarten, and although it is not the purpose of this chapter to compare the effects found in the one and two year analyses, it is reasonable to indicate some of the more outstanding differences between them. The interested reader is referred to Chapter II of this volume for a further development of this issue.

In this context, it can be noted that in the roughly comparable areas of MAT Numbers, measured in the kindergarten battery, and the MAT Mathematics subtest, measured in the first grade battery, there is a reversal of effect found in the one and two year analyses. Thus, the children participating in the Oregon program showed significant math achievement over the NFT group after one year in the program, and they showed a significant deficit in math compared to the NFT group after two years in the program. In addition, the kindergarten measure, Listening for Sounds, and the roughly comparable first grade measure, Word Analysis, show rather different effects in the two studies. On these measures, the Oregon children showed much greater achievement than the NFT groups after one year of FT, but they showed only slightly greater achievement after two years in the program. This is a puzzling shift in effects that should be examined further.

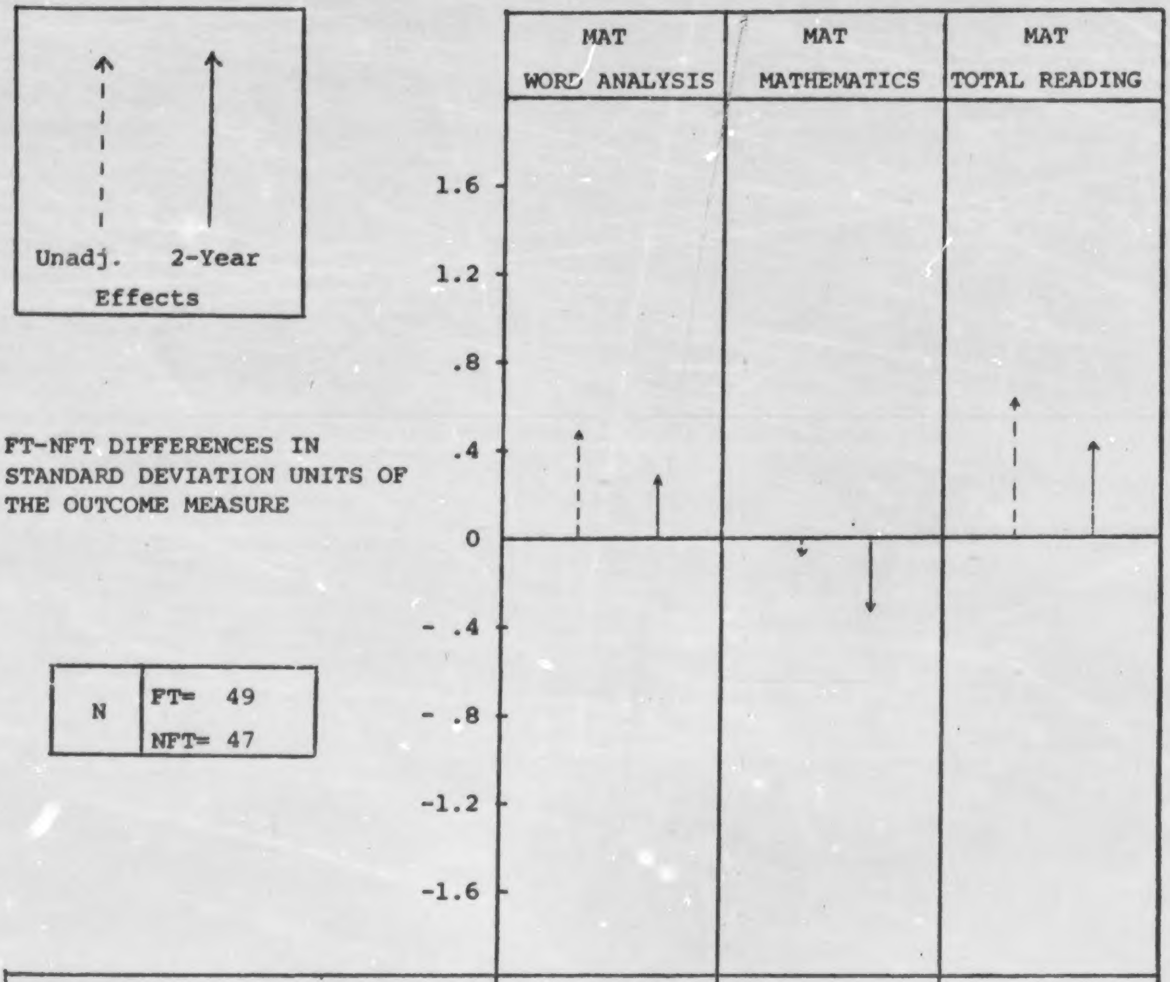
FIGURE VI-19 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
OREGON SITE NC1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.050	- 3.940	8.110
	2-Year	.006	- 5.461	5.648
S.E. = Standard Error of B	Unadjusted	.048	2.424	3.454
	2-Year	.043	2.231	2.936
F = Significance Statistic	Unadjusted	1.086	2.641	5.513*
	2-Year	.038	5.993*	3.702
SD=Raw Score Standard Deviation		.212	10.783	15.639
B/SD = Effect in SD Units (See Graph)	Unadjusted	.236	- .365	.519
	2-Year	.038	- .506	.361

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

FIGURE VI-20 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
OREGON SITE NC2

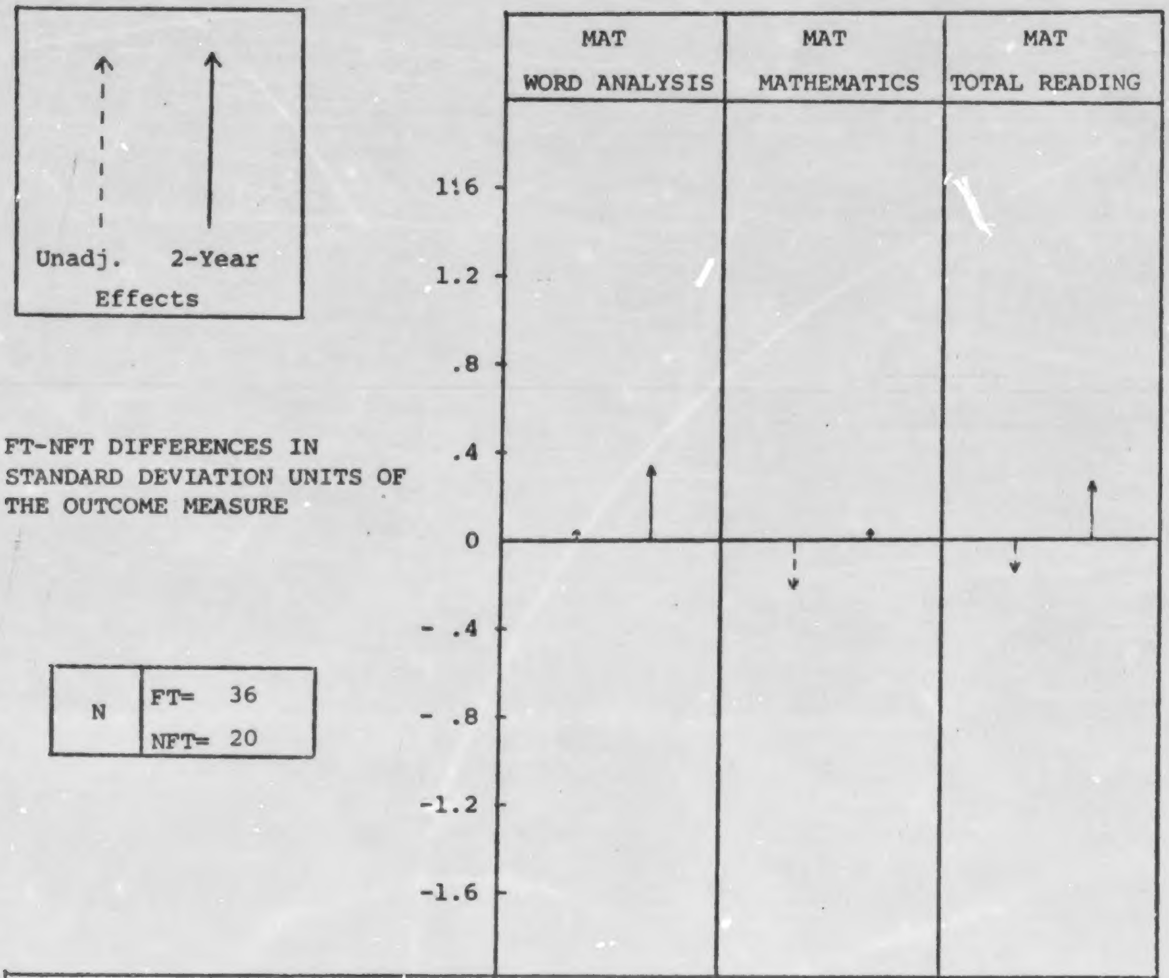


B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.096	- 1.026	10.043
	2-Year	.056	- 3.912	6.775
S.E. = Standard Error of B	Unadjusted	.040	2.350	3.094
	2-Year	.040	2.261	3.013
F = Significance Statistic	Unadjusted	5.760*	.191	10.536*
	2-Year	1.960	1.994	5.058*
SD=Raw Score Standard Deviation		.201	11.462	15.899
B/SD = Effect in SD Units (See Graph)	Unadjusted	.478	- .090	.632
	2-Year	.279	- .341	.426

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-21 END OF FIRST GRADE FT/NFT CONTRASTS

TRACED 2-YEAR SAMPLE
OREGON SITE NE



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.008	- 2.800	- 2.367
	2-Year	.078	2.398	4.527
S.E. = Standard Error of B	Unadjusted	.064	3.516	4.824
	2-Year	.061	3.103	3.907
F = Significance Statistic	Unadjusted	.016	.634	.241
	2-Year	1.636	.598	1.343
SD=Raw Score Standard Deviation		.226	12.566	17.177
B/SD = Effect in SD Units (See Graph)	Unadjusted	.035	- .223	- .138
	2-Year	.345	.032	.264

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

6.5 UNIVERSITY OF KANSAS BEHAVIOR ANALYSIS APPROACH

The primary objective of the University of Kansas approach is to facilitate the child's mastery of basic skills, particularly in reading and arithmetic, through the establishment of a "token economy" within the classroom. Based on basic principles of behavior modification, the token exchange system is designed to provide systematic, positive reinforcement for desired behavior. The tokens, which are given as rewards for successful completion of tasks, may later be exchanged for desired activities. Within the "token economy" environment, programmed instructional materials are used to teach basic skills.

6.5.1 Overall 2-Year Effects

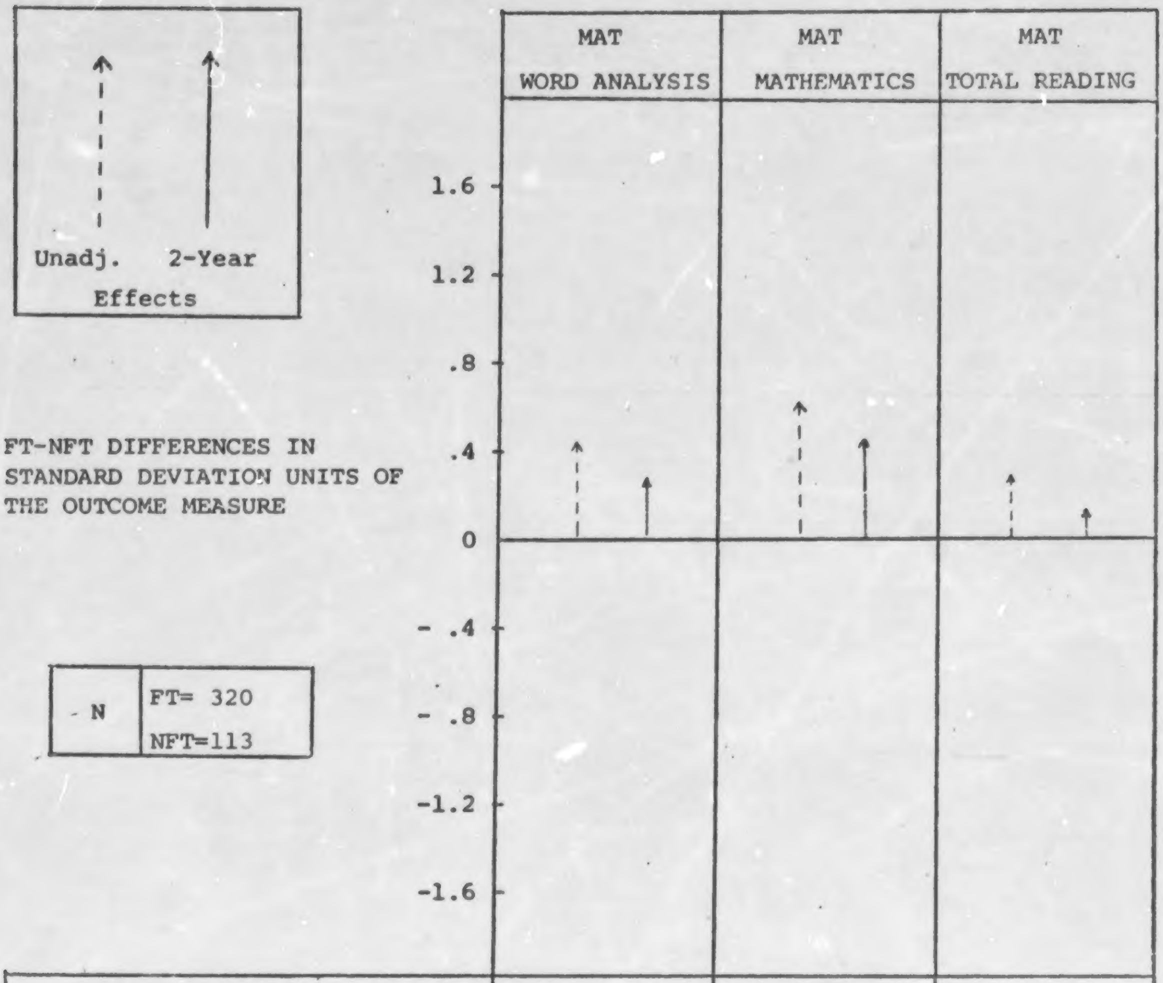
The FT/NFT differences for Kansas shown in Figure VI-22, indicate FT-favoring effects on Word Analysis and Mathematics but no difference on Total Reading. Since the Kansas FT and NFT groups are well matched on background characteristics, the FT-favoring differences at the end of first grade may represent major impacts of the Kansas curriculum. However, the match is not consistent across all sites, nor are the effects consistent across sites. As we shall see, one site has a deviant sample, and effects in this site are quite different from the other Kansas site effects.

6.5.2 Site Descriptions

This analysis of the Kansas program examines data from two types of sites. The family and pupil characteristics are shown in Tables VI-16 through VI-19. The first type includes three large cities, one in the Northeast, one in the North Central, and one in the Southern regions of the country. The children attending the FT and NFT schools in these locations are almost exclusively lower SES Black children. The second type is a rural site in the North Central region in which both FT and NFT groups are predominantly White and the FT children are higher in SES and kindergarten entry level than the NFT children in that site.

The demographic properties of the FT and NFT groups in the three urban sites show an almost identical pattern in all three sites. The ethnic composition in all three is greater than 95% Black in both FT and NFT

FIGURE VI-22 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
KANSAS



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.119	8.272	4.892
	2-Year	.075	6.080	2.009
S.E. = Standard Error of B	Unadjusted	.027	1.443	1.852
	2-Year	.023	1.227	1.559
F = Significance Statistic	Unadjusted	18.989*	32.863*	6.974
	2-Year	11.010*	24.541*	1.660
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	.446	.632	.279
	2-Year	.283	.464	.115

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

TABLE VI-16

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR KANSAS SITE NE

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.07	1.09	60	5.27	.96	15
Father's Education	4.50	1.34	22	4.70	1.34	10
Income	7.21	5.29	61	9.86	5.23	14
Occupation	16.17	1.43	58	15.87	1.36	15
Family Structure						
Mother and Father Present	35%	---	22	67%	---	10
Mother, not Father, Present	60	---	38	26	---	4
Other	5	---	3	7	---	1
TOTAL			63			15
Entry Level						
WRAT	25.68	10.19	109	27.62	10.02	24
PPVT	38.72	11.75	109	38.33	15.48	24
Preschool Experience						
Attended Headstart or other Preschool	98%	---	134	100%	---	1
TOTAL			137			1
Ethnicity						
Black	95%	---	104	97%	---	23
White	1	---	1	3	---	1
Spanish	4	---	4	0	---	0
Other	0	---	0	0	---	0
TOTAL			109			24

TABLE VI-17

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

Kansas Site NC2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.62	.88	61	5.15	.78	34
Father's Education	5.94	.87	18	5.53	.74	15
Income	5.36	5.03	67	7.17	6.34	41
Occupation	15.94	2.05	68	16.29	1.96	41
Family Structure						
Mother and Father Present	25%	-	17	34%	-	14
Mother, not Father, Present	65	-	44	49	-	20
Other	10	-	7	17	-	7
TOTAL			68			41
Entry Level						
WRAT	26.28	9.78	97	28.61	11.76	38
PPVT	39.40	11.57	97	39.82	11.87	38
Preschool Experience						
Attended Headstart or other Preschool	56%	-	62	86%	-	36
TOTAL			110			42
Ethnicity						
Black	100%	-	97	100%	-	38
White	0	-	0	0	-	0
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			97			38

TABLE VI-18

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

Kansas Site S

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.67	.87	64	5.83	.86	35
Father's Education	5.60	.87	25	5.30	1.08	20
Income	6.33	5.83	73	11.00	7.86	39
Occupation	15.84	1.95	71	15.50	1.80	40
Family Structure						
Mother and Father Present	31%	-	23	48%	-	20
Mother, not Father, Present	53	-	39	37	-	15
Other	16	-	12	15	-	6
TOTAL			74			41
Entry Level						
WRAT	23.93	11.60	68	20.83	9.27	29
PPVT	36.30	12.95	68	41.66	13.00	29
Preschool Experience						
Attended Headstart or other Preschool	73%	-	47	74%	-	20
TOTAL			64			27
Ethnicity						
Black	100%	-	68	100%	-	29
White	0	-	0	0	-	0
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			68			29

TABLE VI-19

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

Kansas Site NC1

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.22	1.46	55	4.76	1.38	38
Father's Education	5.02	1.50	50	4.23	1.88	35
Income	9.91	8.00	55	7.40	7.78	42
Occupation	14.93	2.14	54	15.00	2.57	41
Family Structure						
Mother and Father Present	85%	-	48	76%	-	32
Mother, not Father, Present	9	-	5	10	-	4
Other	6	-	3	14	-	6
TOTAL			56			42
Entry Level						
WRAT	33.57	12.39	58	24.68	13.30	40
PPVT	43.84	11.23	58	42.22	15.87	40
Preschool Experience						
Attended Headstart or other Preschool	66%	-	39	55%	-	23
TOTAL			59			42
Ethnicity						
Black	30%	-	17	36%	-	23
White	70	-	41	64	-	26
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			58			40

groups. Preschool attendance is high in all groups in all sites (mean FT = 76%; mean NFT = 86%), and Title I expenditures are high as well. The FT/NFT differences on background variables are almost identical in all sites: FT groups are somewhat lower in total family income, and there is a smaller percentage of intact families in the FT groups in each site. It is likely that father absence accounts for the difference in income since there are fewer wage earners in these families. These characteristics suggest that the social status differences between the FT and NFT groups are not as great as the income differential implies. However, the difference in family structure does represent a substantial difference between the groups.

Kansas' one rural site (NC1, Table VI-19) is a small midwestern town. Here the sample has predominantly intact families (FT = 86%; NFT =76%), and fewer children attend preschool (FT = 66%; NFT = 55%) compared to the urban sites. Despite these demographic similarities, the FT children in this site rank consistently higher on all SES indicators than the NFT group. These status differences are also reflected in sharp differences between FT and NFT on initial entry level. The FT group scores almost a full standard deviation higher than the NFT on the fall kindergarten WRAT.

It seems likely that this site accounts for one of the issues raised in the earlier studies involving the Kansas model. In the analysis of the kindergarten data for this Sponsor reported in Volume 1A of "Education as Experimentation", pretest differences in favor of the FT group were identified. Such differences pose severe problems for analysis of covariance and consequently threatened the interpretability of the analyses of the Kansas data. Two hypotheses were generated to account for the differences. The first was that the Sponsor had achieved a very early impact and that the differences were in fact treatment effects. The second hypothesis was that the two groups were truly different upon entry into the program and that unknown selection procedures, which confound treatment effects, were responsible for the assignment of the higher achieving, higher SES children to FT and the lower achieving children to NFT in the communities in which this Sponsor is working. Although the data available at that time tended to support the latter hypothesis, the issue was not resolved.

The present data cast some new light on this issue. The overall analysis of the Kansas data, in which all children from all sites are aggregated to the Sponsor level, reveals the same FT/NFT differences on kindergarten pretests found in the comparable analysis run for the earlier report. However, these data can be disaggregated to the site level, and here it is possible to see that the source of the difference is to be found primarily in site NC1, the small rural town. Thus, the overall WRAT entry means for FT and NFT respectively are 27 and 24, representing a significant difference in favor of the FT group. However, site NC1 shows the only significant difference on mean entry level scores of the four sites included in the analysis (FT = 33.57; NFT = 24.68); when the children in this site are removed from the overall mean, the difference between the remaining FT and NFT means reduces to 25 and 24, respectively, which is not significant. It is likely, therefore, that the SES differences of the children in this site are the source of the FT/NFT differences in the overall analysis.

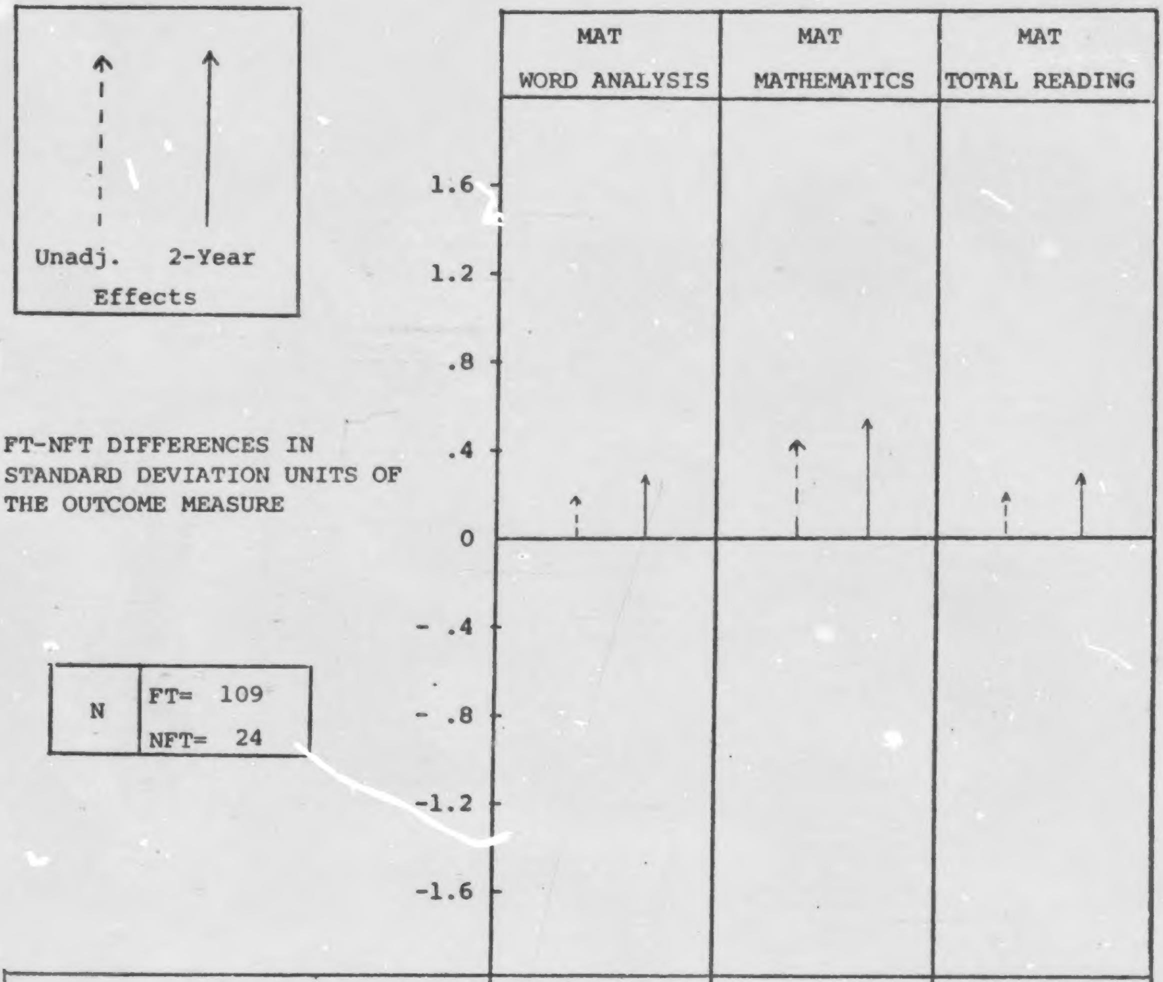
6.5.3 By Site 2-Year Effects

In the three urban sites involving well matched, predominantly Black, lower SES groups, the FT children are ahead of NFT on all outcome measures and sometimes by substantial margins (See Figures VI-23 through VI-26). Kansas appears to be consistently effective across all of these sites after two years, as it was after one year.

In the remaining rural site, NC1 where the FT group contains a higher percentage of White children and is higher on SES than NFT, the program does not appear to be as effective. The NFT group has higher adjusted Total Reading and Word Analysis scores, and there is no difference between the groups on Mathematics. Since the FT group in this site was approximately .7 standard deviations ahead of the NFT on the WRAT at entry to kindergarten, these findings indicate that the NFT group has caught up to the higher entry level, higher SES FT group.

The differences among the sites suggest that the effectiveness of the program may depend on the site (urban or rural), on the ethnic background of the children, or on both. In these data, however, the effects,

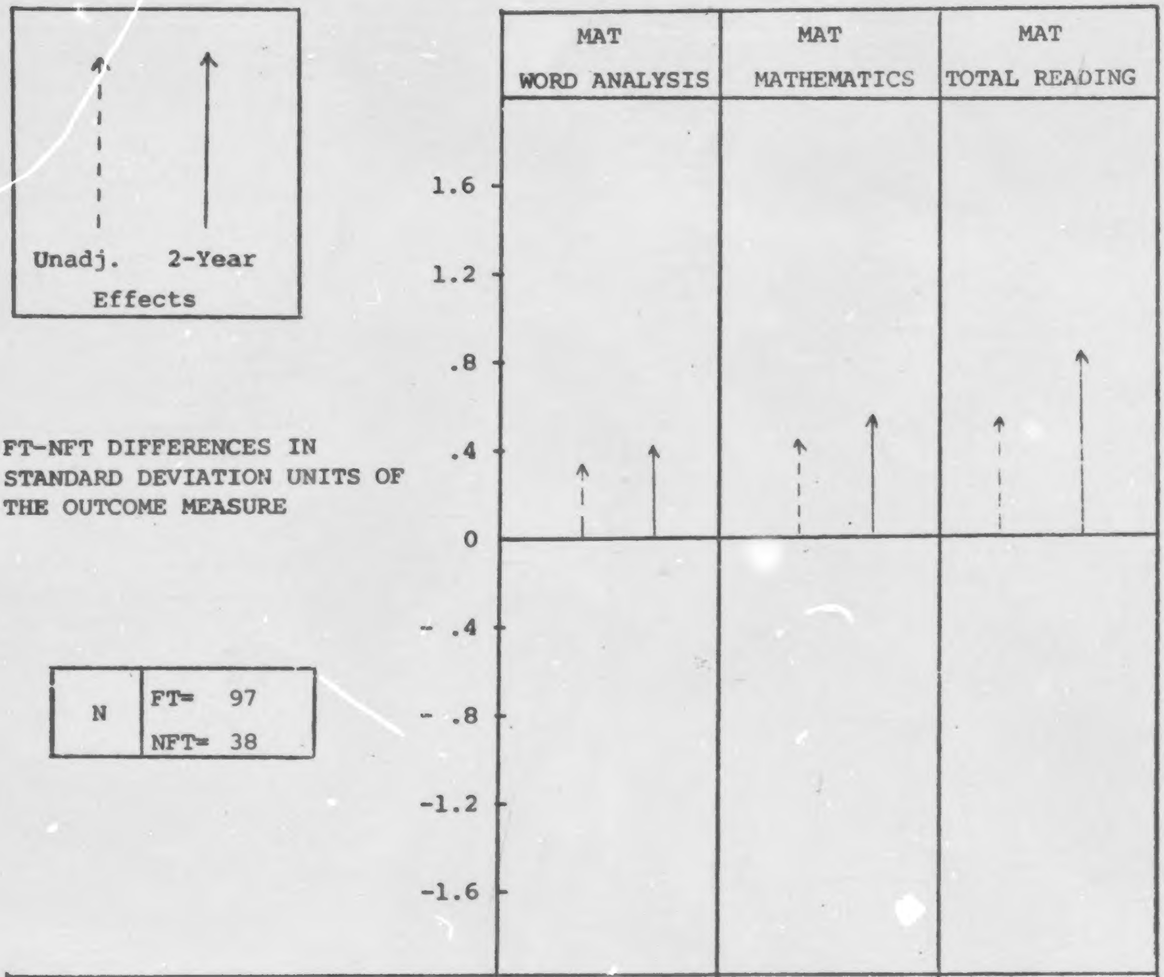
FIGURE VI-23 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
KANSAS SITE NE



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.039	4.600	2.857
	2-Year	.062	5.687	4.523
S.E. = Standard Error of B	Unadjusted	.051	2.302	3.408
	2-Year	.044	1.958	2.879
F = Significance Statistic	Unadjusted	.585	3.992*	.702
	2-Year	1.985	8.433*	2.468
SD=Raw Score Standard Deviation		.228	10.324	15.095
B/SD = Effect in SD Units (See Graph)	Unadjusted	.171	.446	.189
	2-Year	.272	.551	.300

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

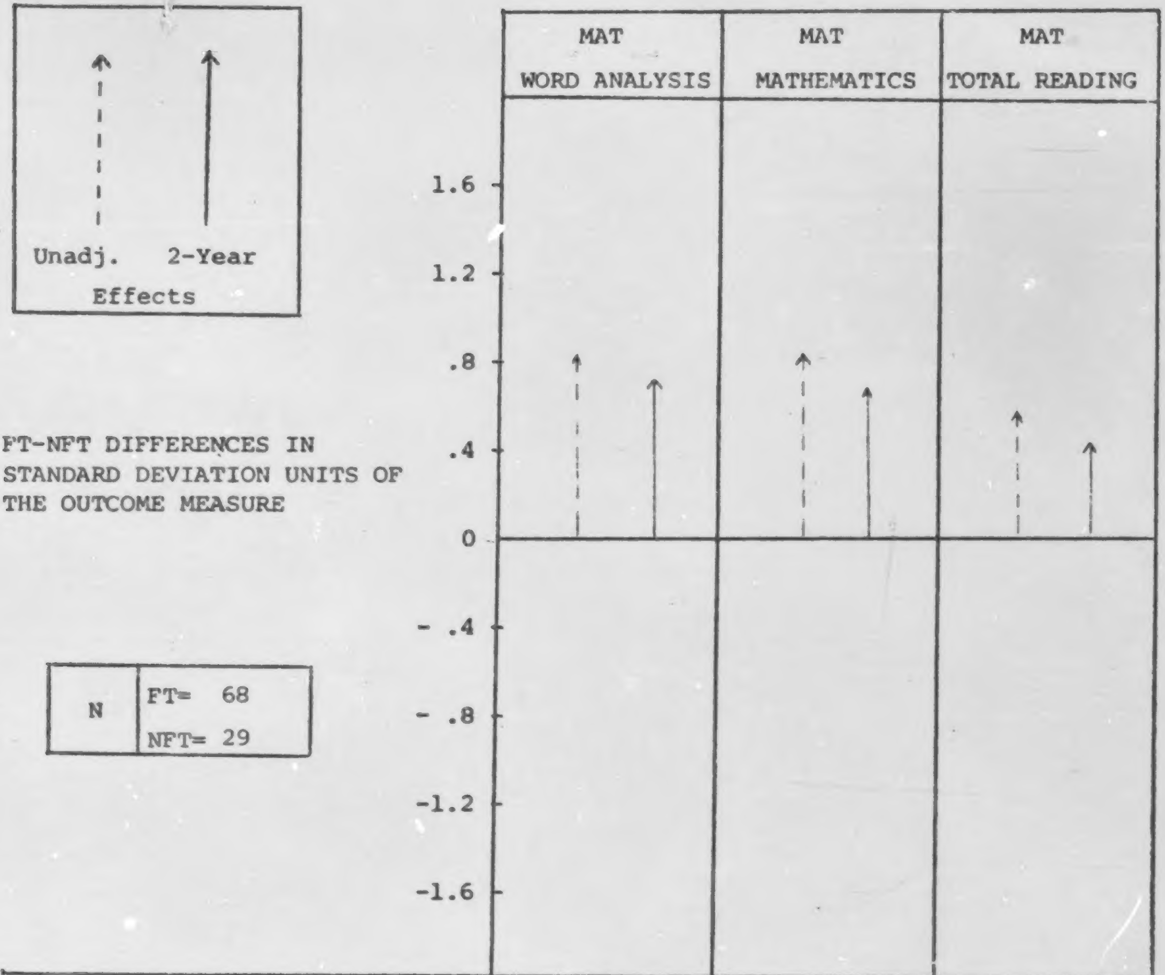
FIGURE VI-24 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
KANSAS SITE NC2



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.077	5.004	3.228
	2-Year	.100	6.182	5.077
S.E. = Standard Error of B	Unadjusted	.045	2.133	3.063
	2-Year	.035	1.794	2.446
F = Significance Statistic	Unadjusted	2.928	5.504*	1.111
	2-Year	8.162*	11.875*	4.310*
SD=Raw Score Standard Deviation		.238	11.332	6.010
B/SD = Effect in SD Units (See Graph)	Unadjusted	.324	.442	.537
	2-Year	.420	.546	.845

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

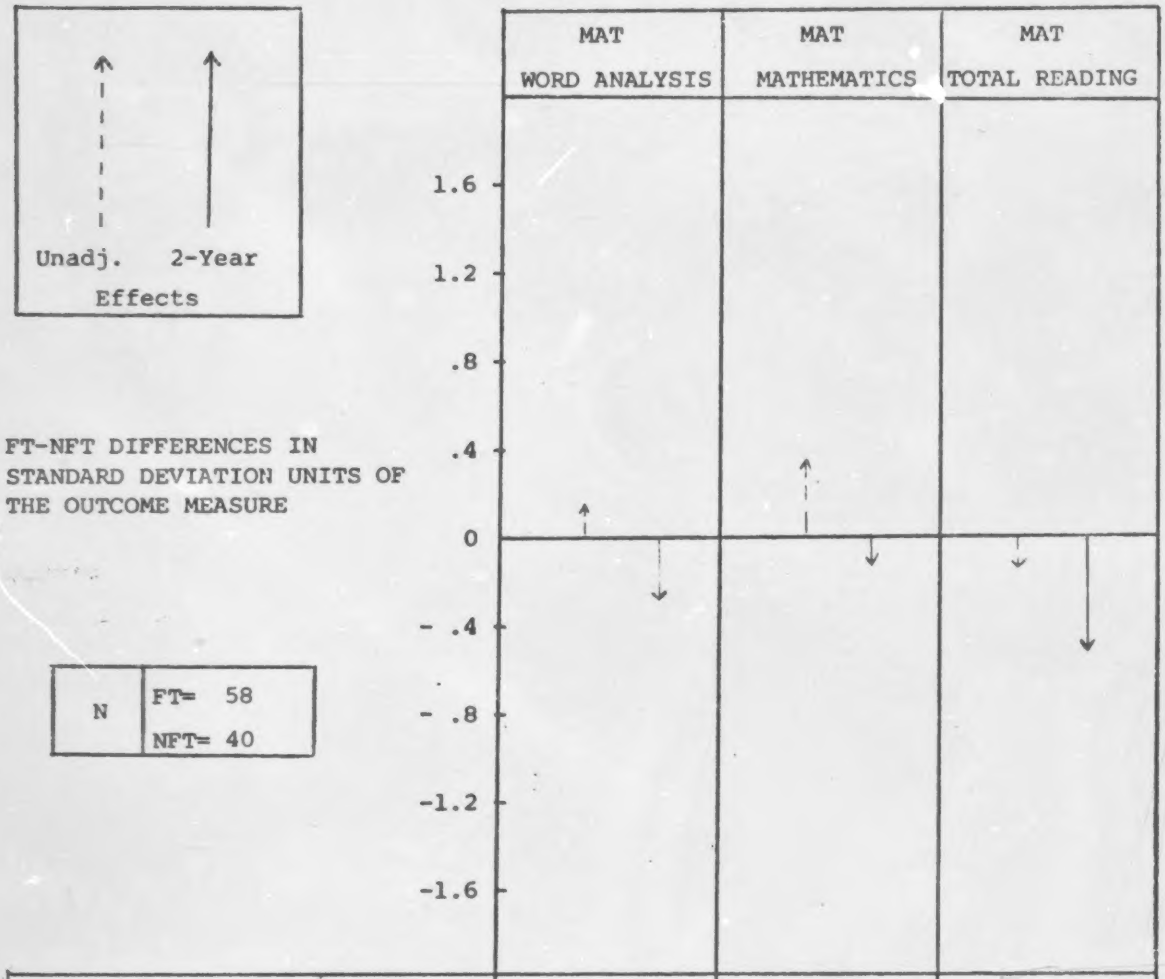
FIGURE VI-25 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
KANSAS SITE S



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.231	11.746	9.776
	2-Year	.199	9.498	7.213
S.E. = Standard Error of B	Unadjusted	.057	2.850	3.597
	2-Year	.048	2.294	2.900
F = Significance Statistic	Unadjusted	16.427*	16.983*	7.388*
	2-Year	17.189*	17.140*	6.185*
SD=Raw Score Standard Deviation		.278	13.877	16.750
B/SD = Effect in SD Units (See Graph)	Unadjusted	.831	.846	.584
	2-Year	.716	.684	.431

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-26 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
KANSAS SITE NC1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.033	4.384	- 2.453
	2-Year	- .072	- 1.844	- 9.436
S.E. = Standard Error of B	Unadjusted	.051	2.608	3.615
	2-Year	.042	1.890	3.064
F = Significance Statistic	Unadjusted	.419	2.826	.461
	2-Year	2.938	.953	9.486*
SD=Raw Score Standard Deviation		.246	12.808	17.540
B/SD = Effect in SD Units (See Graph)	Unadjusted	.134	.342	- .140
	2-Year	- .293	- .144	- .538

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

if any, of site and ethnic background are confounded, and no conclusions about their separate effects can be drawn. The Kansas program does appear to produce substantial academic gains with Black children in the urban sites and to show no meaningful effects in the rural site with White children. These issues clearly need to be explored before a full understanding of the effects of the Kansas program can be had.

6.6 HIGH/SCOPE FOUNDATION, COGNITIVELY ORIENTED CURRICULUM

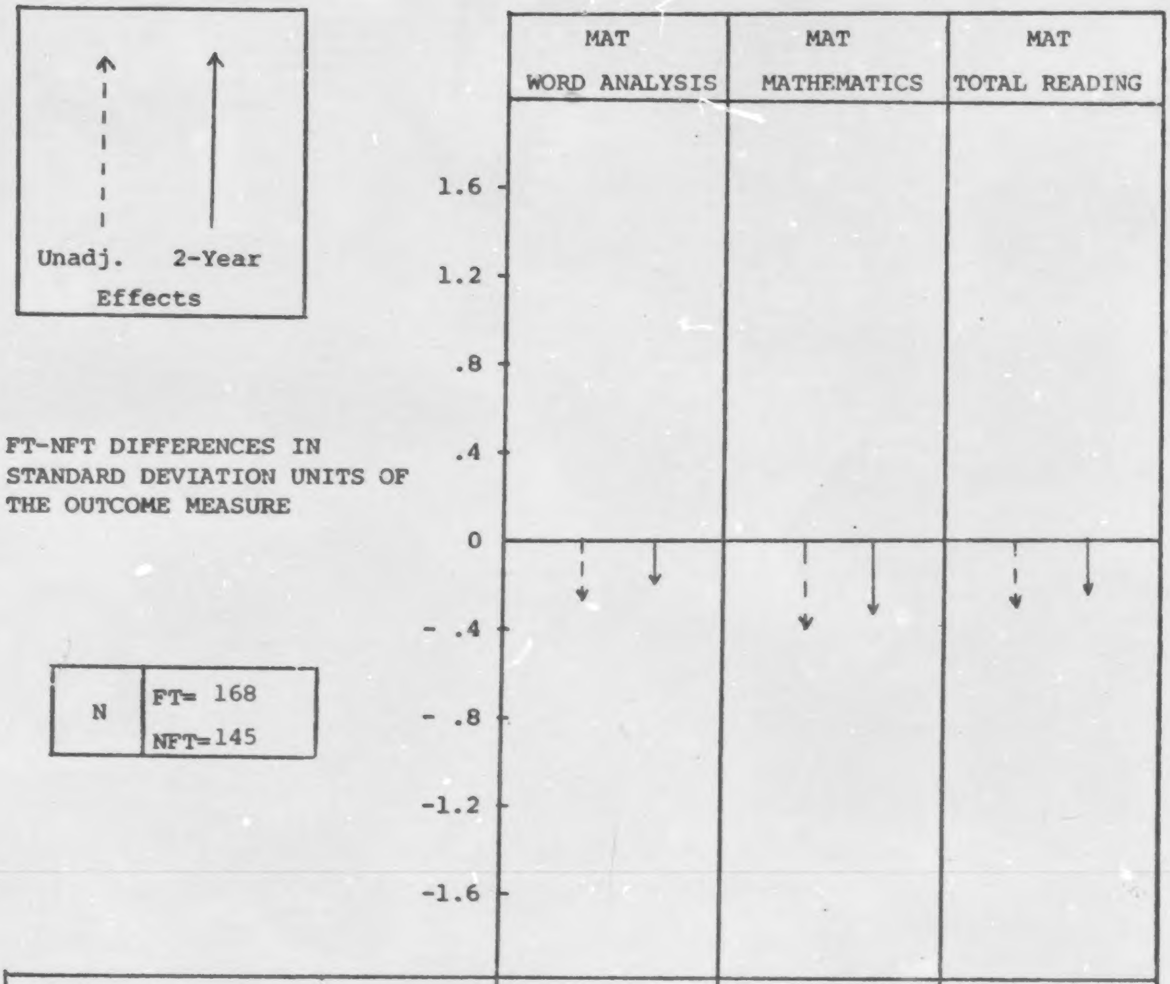
The High/Scope classroom environment may be described as "open", with an emphasis on active exploration, manipulation, and discovery. Within this open frameowrk, the instructional approach is systematic and planned. The cognitively oriented curriculum is highly Piagetian. The ultimate goal is to develop in children the thinking skills they will need throughout their school years and adult lives.

6.6.1 Overall 2-Year Effects

The end of first grade FT/NFT differences for High/Scope, shown in Figure VI-27, indicate a small difference in favor of NFT in Mathematics and no particular differences on other outcomes. Since the FT group is disadvantaged relative to NFT on several SES indicators, these small differences may represent a reversal of the often-observed cumulative deficit in academic performance. Such a reversal may well have begun in preschool since a larger number of FT children (86%) attended preschool than NFT children (56%). This might account for the equality of pre-kindergarten scores despite SES differences. After two years in Follow Through this equality still exists, suggesting that High/Scope is able to continue the potential gains of preschool. This is variable across sites, however, depending upon the kind of children participating in the program.

The High/Scope FT and NFT groups show no substantial differences on entry test scores; however, in most cases, the FT group has far fewer intact families, lower income, and more minority children than NFT. This inconsistency of background and entry test scores may be accounted for by the differences between the percentages of children who attended preschool in the two groups; 84% of the FT group and 56% of the NFT group attended preschool. Thus, the adjustment of posttest scores accomplished by considering pretest scores may remove the benefit that the FT children gained in preschool, a benefit that they may have maintained over the kindergarten and first grade years. Since the differences between FT and NFT on posttests are small, even where NFT is significantly ahead, a small change in the adjustment in the FT direction would present a more favorable picture.

FIGURE VI-27 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH / SCOPE



B = Sponsor Effect	Unadjusted	- .072	- 5.231	- 5.424
(FT-NFT Differences)	2-Year	- .053	- 4.411	- 4.331
S.E. = Standard	Unadjusted	.029	1.283	1.627
Error of B	2-Year	.024	1.101	1.364
F = Significance	Unadjusted	6.237*	16.613*	11.112*
Statistic	2-Year	4.697*	16.040*	10.009*
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD	Unadjusted	- .269	- .400	- .309
Units (See Graph)	2-Year	- .198	- .337	- .247

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

6.6.2 Site Descriptions

Included in this analysis of High/Scope effects are six geographically diverse sites with diverse characteristics. In four of the sites the FT/NFT groups are fairly well matched on SES and entering achievement; in two, they are severely mismatched. The pupil and family characteristics for the sites are shown in Tables VI-20 through VI-25.

The four sites in which High/Scope's FT and NFT groups are closely matched are geographically scattered; two are large cities, one in the Northeast (NE, Table VI-20) and the other in the North Central Region (NC, Table VI-21). The other two sites are in the West, one a middle-sized city (W1, Table VI-22); the other a rural town (W2, Table VI-23). In all these sites, the FT and NFT groups are very similar on family structure characteristics and on all available SES indicators except income level, which is lower in the FT groups. The FT and NFT children in the Northern sites are nearly all Black; there is a small percentage (4%) of Spanish-surnamed children at the Northeastern site. Two of the Western sites are predominantly Spanish-surname and White; site W3 is predominantly Black. None of the entry level scores at these sites show significant differences.

The two High/Scope sites which show a severe mismatch between FT and NFT groups show no difference on parent's level of education, but the groups are different on all other SES indicators. (See S, Table VI-24; W3, Table VI-25). The income level in FT families is lower than in NFT families, and the FT group has considerably fewer intact families. The FT group in site S is predominantly Black while the majority of NFT children are White. While the groups do not differ much on entry level WRAT, the NFT group PPVT scores are much higher then FT's.

Site W3 displays a similar disadvantage in the FT group. However the mismatch appears to be more severe. The FT group is predominantly Black, while the NFT group is a mixture slightly over half White, slightly under half children of various minorities. The FT group is lower than NFT on all SES indicators except the level of mother's education and has considerably fewer intact families. The FT groups also shows a substantial disadvantage on both the kindergarten entry WRAT and PPVT.

TABLE VI-20

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

High/Scope Site NE

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.65	.85	34	5.50	.54	8
Father's Education	5.67	1.03	18	missing		
Income	9.55	7.30	38	missing		
Occupation	15.71	1.80	38	missing		
Family Structure						
Mother and Father Present	42%	-	17	missing		
Mother, not Father, Present	41	-	16	missing		
Other	17	-	7	missing		
TOTAL			40	missing		
Entry Level						
WRAT	29.06	11.46	35	32.76	12.54	17
PPVT	38.54	11.85	35	39.71	9.55	17
Preschool Experience						
Attended Headstart or other Preschool	83%	-	34	94%	-	15
TOTAL			41			16
Ethnicity						
Black	100%	-	35	96%	-	16
White	0	-	0	0	-	0
Spanish	0	-	0	4	-	1
Other	0	-	0	0	-	0
TOTAL			35			17

TABLE VI-21

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

High/Scope Site NC

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.34	.63	35	5.32	1.10	31
Father's Education	4.86	1.24	15	4.83	1.19	12
Income	8.05	6.35	37	9.10	6.33	30
Occupation	missing			missing		
Family Structure						
Mother and Father Present	46%	-	17	47%	-	15
Mother, not Father, Present	45	-	17	47	-	15
Other	9	-	3	6	-	1
TOTAL			37			31
Entry Level						
WRAT	29.37	11.42	40	28.45	10.32	33
PPVT	34.82	10.47	40	34.12	13.93	33
Preschool Experience						
Attended Headstart or other Preschool	82%	-	33	27%	-	3
TOTAL			40			11
Ethnicity						
Black	100%	-	40	100%	-	33
White	0	-	0	0	-	0
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			40			33

TABLE VI-22

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

High/Scope Site W1

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	4.74	1.71	34	4.69	1.08	16
Father's Education	4.68	1.44	28	4.29	1.54	14
Income	10.08	7.08	36	10.83	6.25	18
Occupation	15.50	1.81	34	15.71	2.47	17
Family Structure						
Mother and Father Present	72%	-	26	66%	-	12
Mother, not Father, Present	22	-	8	17	-	3
Other	6	-	2	17	-	3
TOTAL			36			18
Entry Level						
WRAT	31.19	10.06	36	34.22	12.46	18
PPVT	52.11	8.58	36	52.50	11.08	18
Preschool Experience						
Attended Headstart or other Preschool	75%	-	27	67%	-	12
TOTAL			36			18
Ethnicity						
Black	0%	-	0	5%	-	1
White	33	-	12	28	-	5
Spanish	67	-	24	67	-	12
Other	0	-	0	0	-	0
TOTAL			36			18

TABLE VI-23

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

High/Scope Site W2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.08	1.28	39	4.82	1.34	40
Father's Education	4.83	1.13	18	5.32	1.19	28
Income	5.66	5.20	38	9.43	6.74	42
Occupation	16.08	1.69	37	15.47	1.66	38
Family Structure						
Mother and Father Present	41%	-	16	67%	-	28
Mother, not Father, Present	54	-	21	29	-	12
Other	5	-	2	4	-	2
TOTAL			39			42
Entry Level						
WRAT	26.03	9.72	34	28.00	10.00	40
PPVT	45.82	10.06	34	43.05	12.24	40
Preschool Experience						
Attended Headstart or other Preschool	94%	-	34	43%	-	22
TOTAL			36			51
Ethnicity						
Black	29%	-	10	12%	-	5
White	17	-	6	20	-	8
Spanish	54	-	18	63	-	25
Other	0	-	0	5	-	2
TOTAL			34			40

TABLE VI-24

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

High/Scope Site S

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.26	.93	46	5.16	1.11	59
Father's Education	4.96	1.23	28	5.35	1.31	49
Income	6.10	4.69	48	10.34	7.10	60
Occupation	missing			missing		
Family Structure						
Mother and Father Present	48%	-	23	71%	-	42
Mother, not Father, Present	45	-	22	23	-	14
Other	7	-	3	6	-	4
TOTAL			48			60
Entry Level						
WRAT	19.71	11.07	52	21.69	10.[illegible]9	60
PPVT	34.62	10.51	52	41.56	13.53	60
Preschool Experience						
Attended Headstart or other Preschool	91%	-	34	missing		
TOTAL			37	missing		
Ethnicity						
Black	84%	-	43	26	-	16
White	15	-	8	74	-	44
Spanish	1	-	1	0	-	0
Other	0	-	0	0	-	0
TOTAL			52			60

TABLE VI-25

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

High/Scope Site W3

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.86	.89	58	5.90	1.45	29
Father's Education	5.89	1.13	18	6.73	.99	22
Income	7.34	6.39	58	14.64	7.05	28
Occupation	16.10	1.91	57	14.57	2.45	30
Family Structure						
Mother and Father Present	26%	-	15	67%	-	20
Mother, not Father, Present	67	-	39	26	-	8
Other	7	-	4	7	-	2
TOTAL			58			30
Entry Level						
WRAT	29.11	12.45	43	34.97	13.75	34
PPVT	44.07	12.29	43	50.73	13.32	34
Preschool Experience						
Attended Headstart or other Preschool	80%	-		missing		
TOTAL			41	missing		
Ethnicity						
Black	70%	-	30	15%	-	5
White	21	-	9	59	-	20
Spanish	0	-	0	3	-	1
Other	9	-	4	23	-	8
TOTAL			43			34

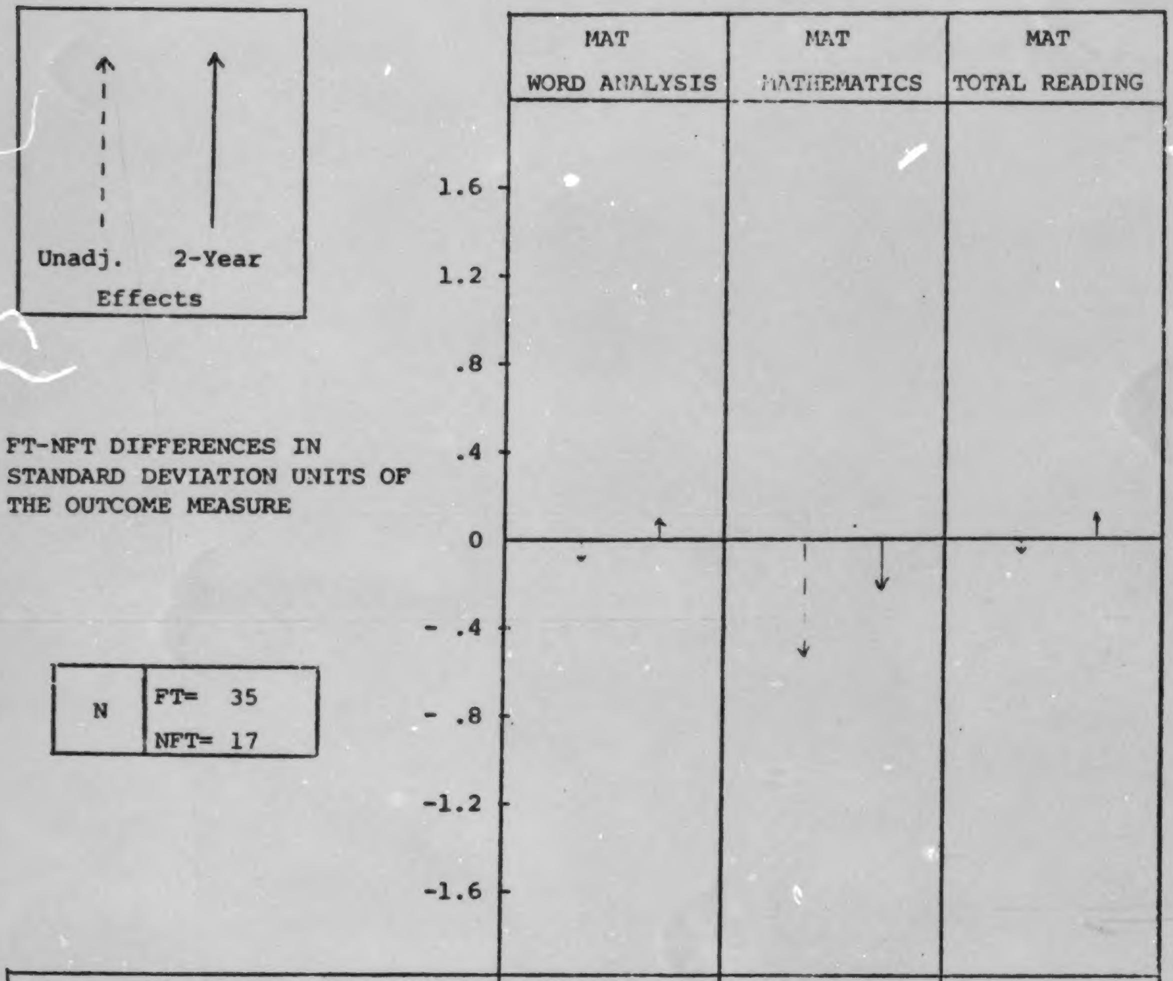
6.6.3 By Site 2-Year Effects

The site level FT/NFT differences for High/Scope on end of first grade measures are shown in Figures VI-28 through VI-33. The pattern of differences across sites is variable. The four sites in which High Scope's FT and NFT groups are closely matched (NE, NC, W1 and W2) show few differences (See Figures VI-28, VI-29, VI-30, and VI-31, respectively). In these four sites, the only substantial differences between the groups are found in site NC where the NFT group gained more on Total Reading and site W2 where the FT group showed greater gain on Word Analysis. There are no obvious factors that account for these variations.

In sites S (Figure VI-32) and W3 (Figure VI-33) the mismatch is severe. Here, the FT groups in general have much lower SES and entry scores. In both sites, the FT groups also do significantly less well than the NFT on most outcome measures. Given the ethnic differences in these sites (FT is predominantly Black, NFT predominantly White and mixed), comparisons are heavily confounded and cannot easily be used to assess the effectiveness of the High/Scope programs.

If we focus on the sites in which reasonable comparisons can be justified, it is clear that High/Scope has been consistently producing the same kinds of FT/NFT contrasts across sites and over time (both one and two year effects). These constrasts show no meaningful differences between the groups in most academic outcomes. The FT children have consistently maintained their relative status over the two years of their FT experience. Since the High/Scope program gives little emphasis to the traditional content of academic tests, it is clear that the FT children are not being hurt (as measured by traditional tests) by this procedure. This is true despite the low performance normally expected from economically disadvantaged children. Future analyses may reveal other aspects of academic growth which High/Scope facilitates, but which are not measured by the FT battery.

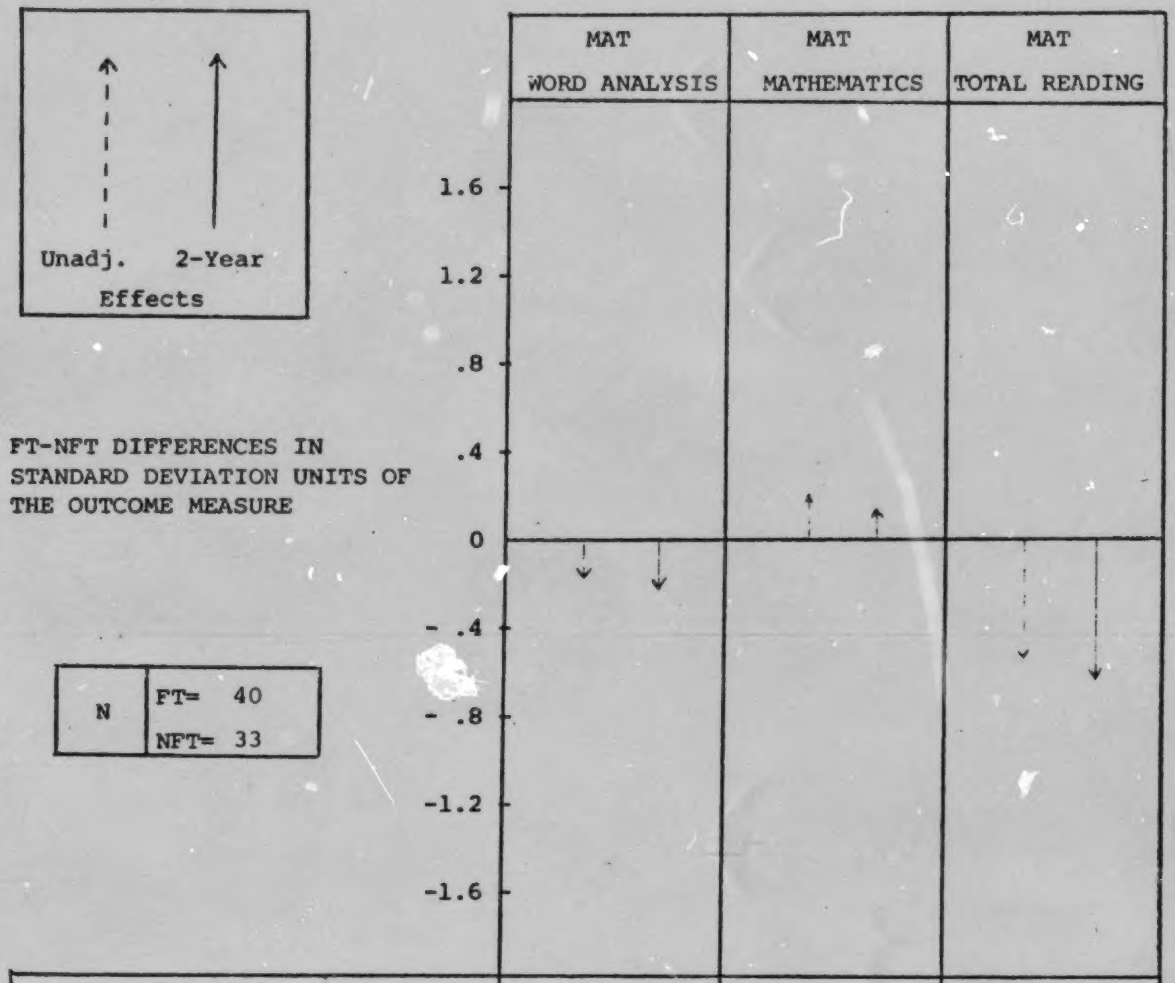
FIGURE VI-28 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH/SCOPE SITE NE



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .019	- 5.264	- .744
	2-Year	.018	- 2.354	1.428
S.E. = Standard Error of B	Unadjusted	.061	2.794	4.069
	2-Year	.054	2.602	3.481
F = Significance Statistic	Unadjusted	.097	3.550	.034
	2-Year	.111	.819	.168
SD=Raw Score Standard Deviation		.203	9.684	13.633
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .094	- .544	- .055
	2-Year	.089	- .243	.105

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

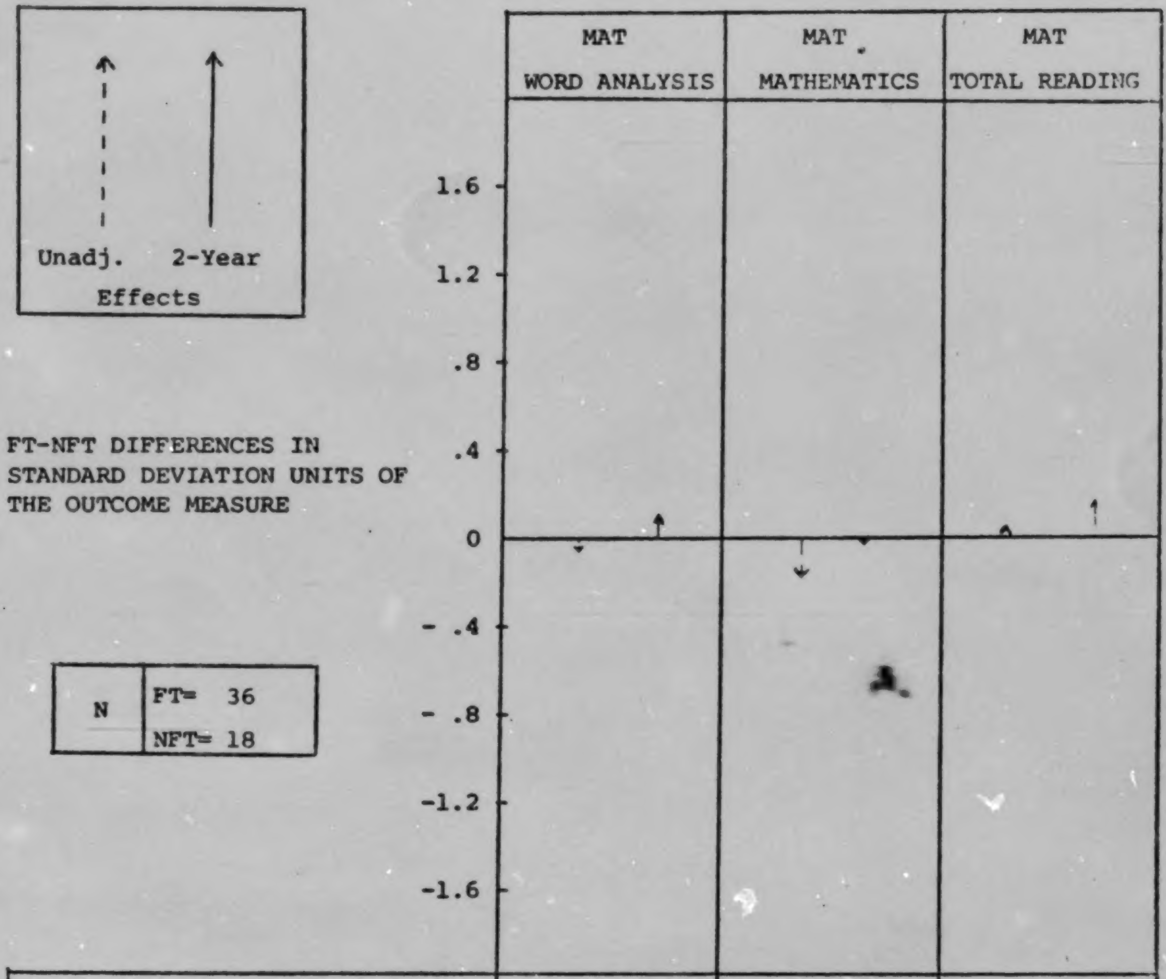
FIGURE VI-29 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH/SCOPE SITE NC



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .038	1.863	- 6.354
	2-Year	- .048	1.291	- 7.230
S.E. = Standard Error of B	Unadjusted	.050	2.281	2.543
	2-Year	.046	1.836	2.078
F = Significance Statistic	Unadjusted	.578	.668	6.245*
	2-Year	1.090	.494	12.103*
SD=Raw Score Standard Deviation		.213	9.679	11.200
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .178	.192	- .567
	2-Year	- .225	.133	- .646

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

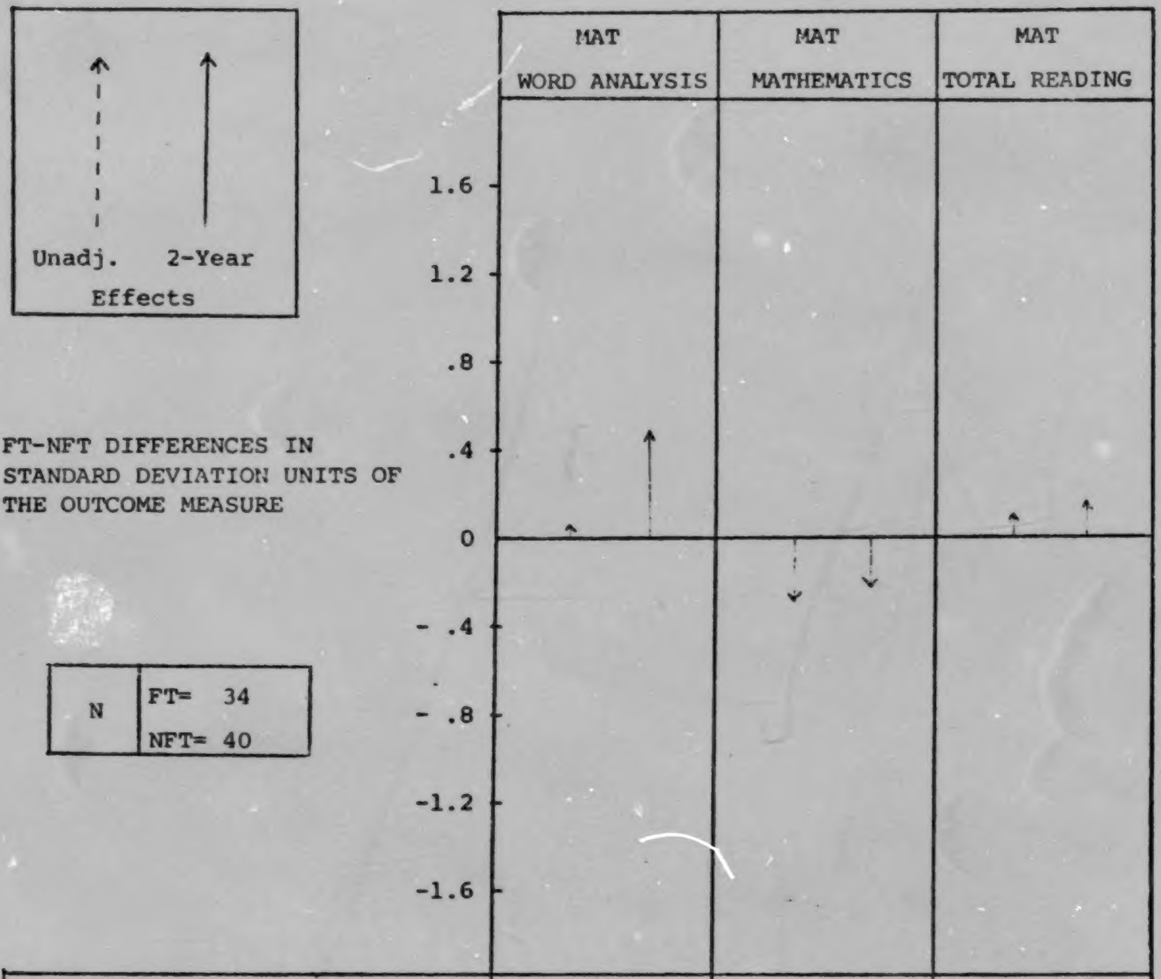
FIGURE VI-30 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH/SCOPE SITE W1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .012	- 1.889	.278
	2-Year	.019	- .139	2.216
S.E. = Standard Error of B	Unadjusted	.058	3.048	4.067
	2-Year	.050	2.482	3.802
F = Significance Statistic	Unadjusted	.043	.384	.005
	2-Year	.144	.003	.340
SD=Raw Score Standard Deviation		.200	10.496	13.956
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .060	- .180	.020
	2-Year	.095	- .013	.159

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

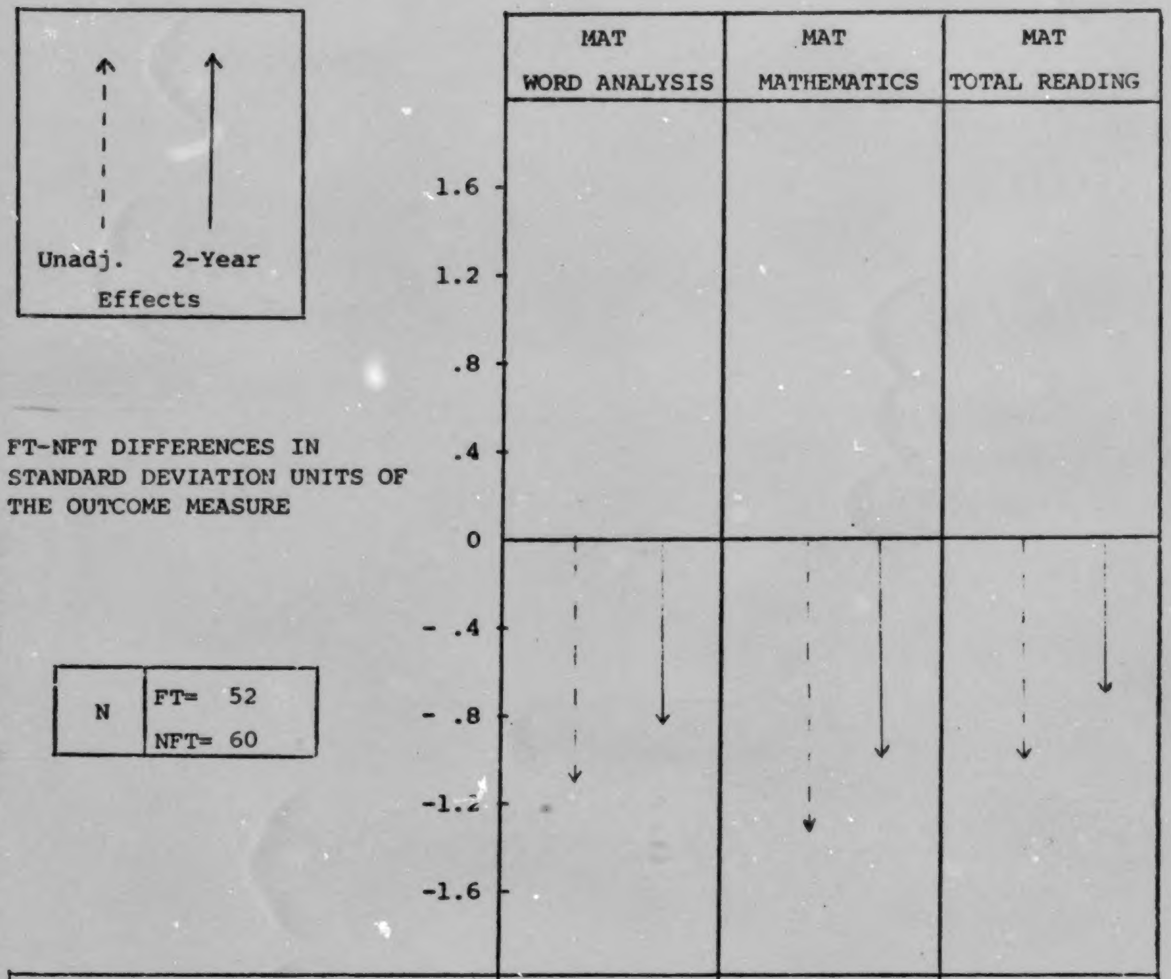
FIGURE VI-31 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH/SCOPE SITE W2



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.008	- 3.228	1.169
	2-Year	.100	- 2.473	2.093
S.E. = Standard Error of B	Unadjusted	.050	2.480	3.018
	2-Year	.047	2.351	2.468
F = Significance Statistic	Unadjusted	.026	1.695	.150
	2-Year	4.528*	.040	.719
SD=Raw Score Standard Deviation		.213	10.681	12.864
B/SD = Effect in SD Units (See Graph)	Unadjusted	.038	- .302	.091
	2-Year	.470	- .232	.163

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

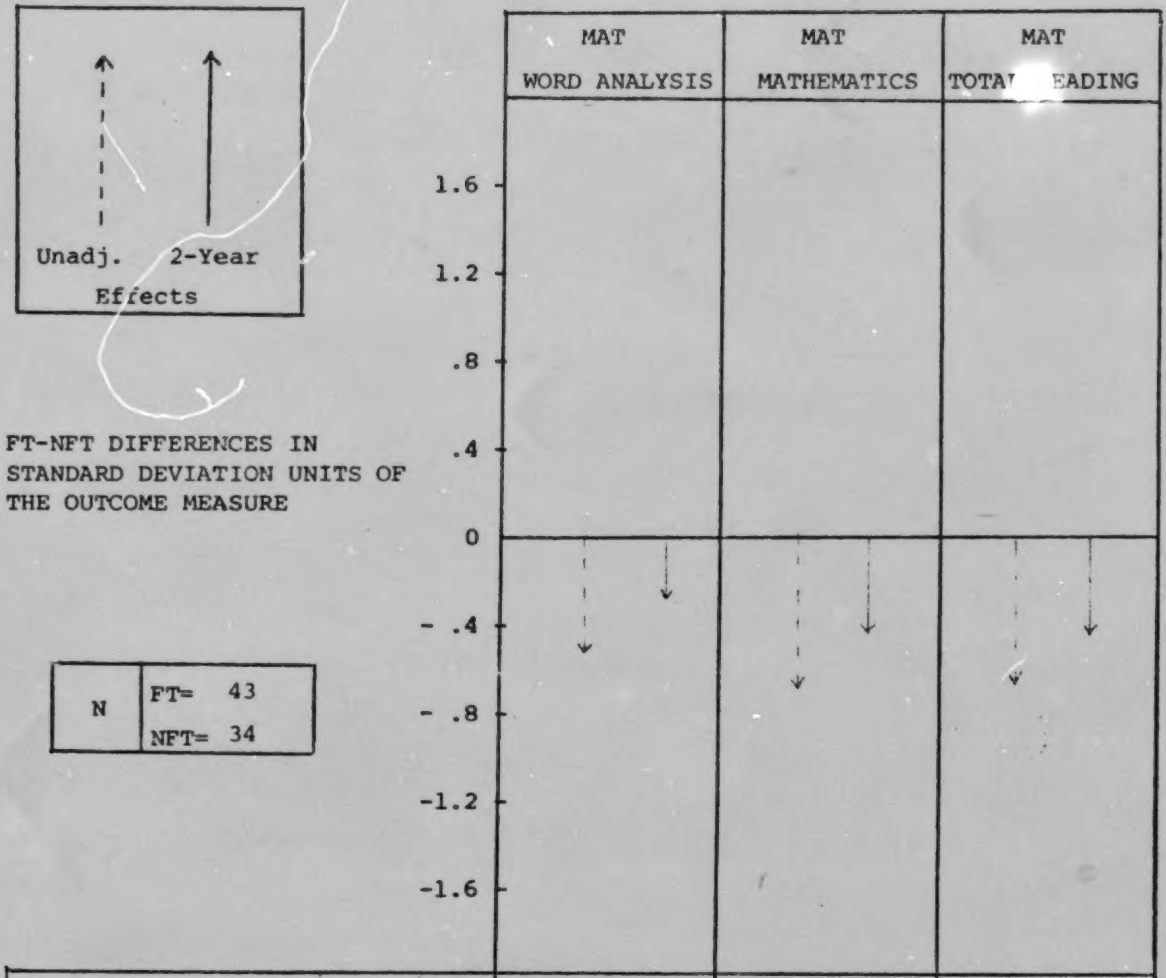
FIGURE VI-32 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH/SCOPE SITE S



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .269	-16.400	-15.969
	2-Year	- .201	-13.204	-10.800
S.E. = Standard Error of B	Unadjusted	.047	2.127	3.012
	2-Year	.039	1.693	2.368
F = Significance Statistic	Unadjusted	32.753*	59.444*	28.111*
	2-Year	26.564*	60.824*	20.803*
SD=Raw Score Standard Deviation		.238	12.109	14.909
B/SD = Effect in SD Units (See Graph)	Unadjusted	- 1.130	- 1.354	- 1.071
	2-Year	- .845	- 1.090	- .724

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-33 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
HIGH/SCOPE SITE W3



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .135	- 8.974	-11.386
	2-Year	- .068	- 5.786	- 7.485
S.E. = Standard Error of B	Unadjusted	.058	2.832	3.686
	2-Year	.046	2.337	3.096
F = Significance Statistic	Unadjusted	5.420*	10.042*	9.542*
	2-Year	2.184	6.131*	5.847*
SD=Raw Score Standard Deviation		.255	12.833	16.650
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .529	- .699	- .684
	2-Year	- .267	- .451	- .450

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

6.7 UNIVERSITY OF FLORIDA PARENT EDUCATIONAL MODEL

The University of Florida program is described as a Parent Education Model. Based on the premise that children's learning takes place as much at home as in school, the major objective is to improve children's school achievement through educating parents to participate directly in the education of their children. While the curriculum is not standardized, it does have a Piagetian orientation.

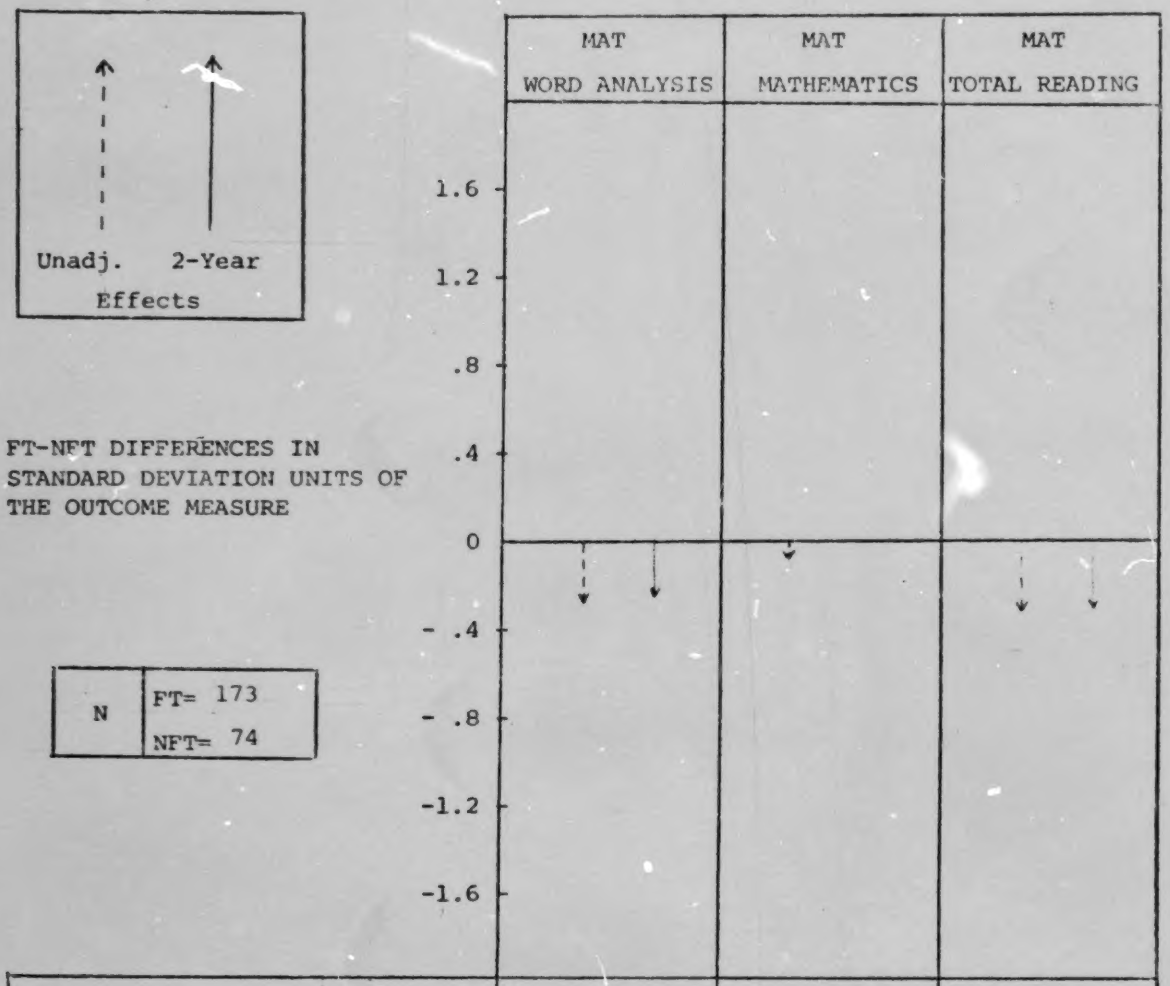
6.7.1 Overall 2-Year Effects

The end of first grade FT/NFT differences for Florida, shown in Figure VI -34, indicate small NFT-favoring differences on Word Analysis and Total Reading. The adjusted difference on Mathematics is not presented because the relationship between this outcome measure and the entry WRAT was different in FT and NFT. These relationships, shown in Figure VI-35, suggest that a relatively large proportion of FT children with low scores on the WRAT scored somewhat higher on the Mathematics posttest than NFT children with comparable pretest scores and that NFT children with higher initial WRAT scores scored somewhat higher on the Mathematics posttest than FT children with comparable pretest scores. These two differences average out, however, producing no net difference on our posttest mean.

Since the FT group is somewhat disadvantaged on background characteristics, having a greater percentage of minority children (67% vs. 30%), fewer intact families, and substantially lower income, the small differences at the end of first grade may represent a favorable event related to the Florida program.

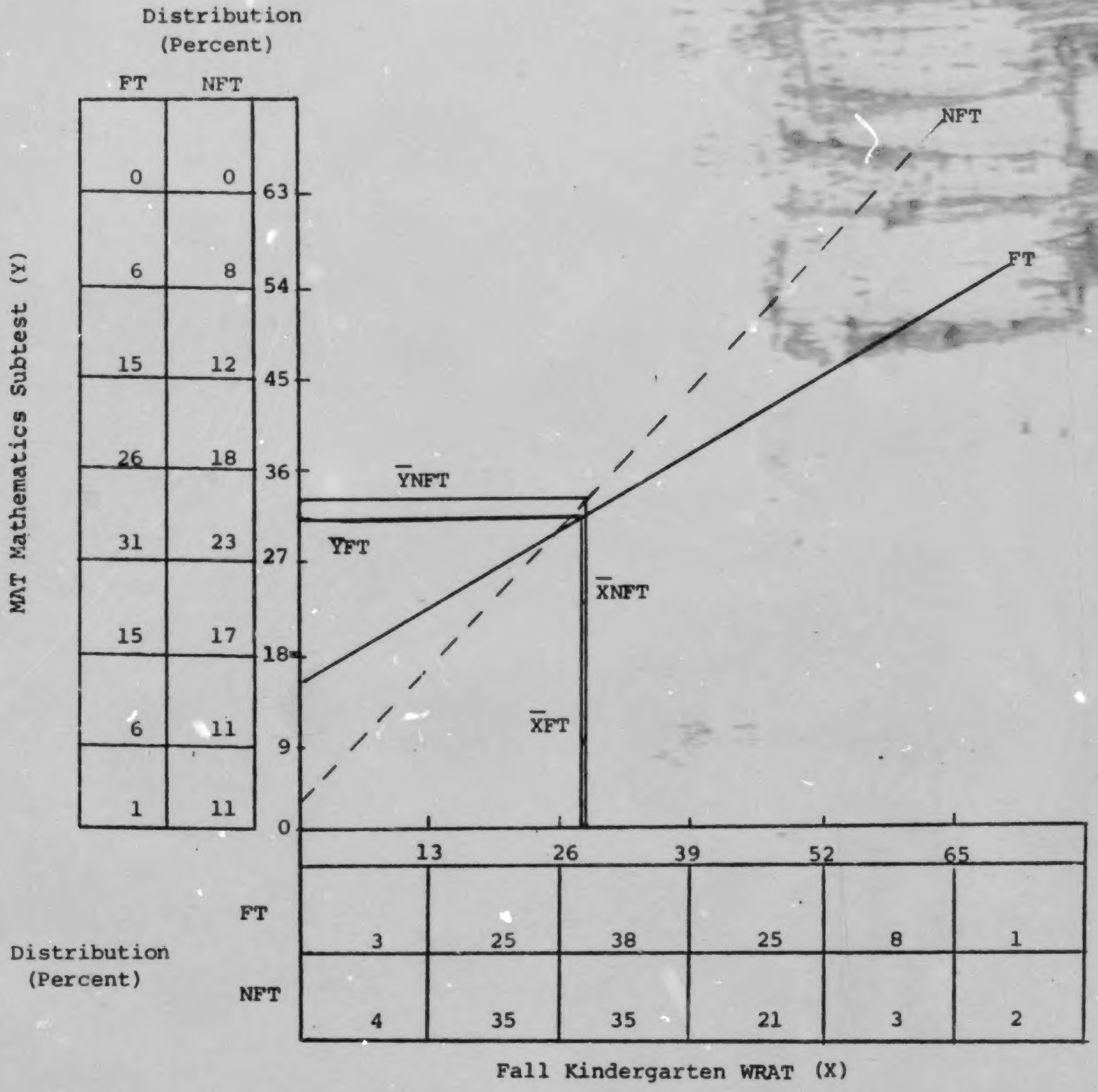
As for several other Sponsors, some of the background characteristics of the Florida FT group suggest that the group is disadvantaged relative to NFT. However, the FT group has entry achievement scores very similar to those of the NFT group. As was the case for these other sponsors, the inconsistency can be explained in terms of the difference in the number of children in the two groups who attended preschool. Whereas 58% of the FT group attended preschool, only 10% of the NFT group had such experience.

FIGURE VI-34 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FLORIDA



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .071	- 1.296	- 5.712
	2-Year	- .067		- 5.410
S.E. = Standard Error of B	Unadjusted	.036	1.707	2.286
	2-Year	.031		1.900
F = Significance Statistic	Unadjusted	3.805*	.576	6.241*
	2-Year	4.705*		8.105*
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .267	- .099	- .326
	2-Year	- .250		- .309

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq .05$



	FT	NFT
$\bar{X}$	27.93	28.35
$\bar{Y}$	31.31	32.61
b	.43	.74
SE (b)	.06	.09
n	173	74

Figure VI-35

Heterogeneous Regressions:
Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT
Sponsor: Florida
Outcome: MAT Mathematics Subtest

Thus, the absence of differences on pretests may be related to the FT children's catching up to NFT as a consequence of their preschool experience. In light of this possibility, the adjustment test may be somewhat unjustified since this measure may reflect the FT group's benefits from preschool and thus may not appropriately represent their status relative to NFT.

6.7.2 Site Descriptions

This evaluation of the Florida program examines data from two types of sites. The family and pupil characteristics are shown in Tables VI-26 through VI-29. The first type consists of three communities in the South (S1, Table VI-26; S2, Table VI-27) and West (W, Table VI-28). In these communities, although the FT and NFT groups differ demographically, they are roughly equal on kindergarten entry level tests. All of these sites display a high degree of social stratification either by ethnic membership or SES, and the FT group comes from the lower socioeconomic strata.

The second type of site is a small North Central city (NC, Table VI-29) in which the FT/NFT children are almost exclusively White. Though both groups have a rather low SES, they are not as poor as the families in the first type of site. The FT children in this site score lower on entry level tests than the NFT children. The lower entry test scores, in the absence of SES differences, suggest that these children are disadvantaged in a manner not directly reflected in SES, but perhaps indicative of an educational disadvantage or learning disability.

The first type of site is represented by two large Southern cities, one in the Southeast (S1), the other in the Southwest (S2), and a small town in the Far West (W).

In the large Southeastern city (S1) 86% of the FT children are Black, compared with 50% in the NFT group. The FT children come from families with significantly lower incomes than those in the NFT group. And the FT group has a higher percentage of children who attended preschool.

In the Southwestern city (S2), virtually all the FT/NFT children are from minority groups. Of the FT children, 63% are Black and 35% Spanish-surnamed; of the NFT children, 42% are Black and 55% Spanish-surnamed. Both groups have very low SES, but the FT group is slightly

TABLE VI-26

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR Florida Site

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.43	.90	54	5.18	.81	33
Father's Education	4.85	.91	27	5.10	.94	21
Income	7.29	5.68	58	10.63	7.32	30
Occupation	15.91	1.84	53	15.42	1.96	31
Family Structure						
Mother and Father Present	44%	-	27	57%	-	21
Mother, not Father, Present	43	-	26	31	-	11
Other	13	-	8	12	-	4
TOTAL			61			36
Entry Level						
WRAT	27.87	13.74	61	26.90	13.00	38
PPVT	44.26	12.67	61	44.37	14.91	38
Preschool Experience						
Attended Headstart or other Preschool	54%	-	35	5%	-	2
TOTAL			64			40
Ethnicity						
Black	86%	-	52	50%	-	19
White	14	-	9	48	-	18
Spanish	0	-	0	0	-	0
Other	0	-	0	2	-	1
TOTAL			61			38

TABLE VI-27

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

Florida Site S2

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.16	1.11	56	4.48	1.03	21
Father's Education	4.81	1.13	26	4.27	1.19	11
Income	6.49	6.43	57	5.22	4.38	23
Occupation	15.93	1.66	60	15.52	1.40	21
Family Structure						
Mother and Father Present	43%	-	26	48%	-	11
Mother, not Father, Present	54	-	32	48	-	11
Other	3	-	2	4		1
TOTAL			60			23
Entry Level						
WRAT	27.76	12.20	102	23.66	9.38	35
PPVT	40.32	13.55	102	37.92	13.85	35
Preschool Experience						
Attended Headstart or other Preschool	99%	-	81	missing		
TOTAL			82	missing		
Ethnicity						
Black	63%	-	64	42%	-	15
White	1	-	1	3	-	1
Spanish	35	-	36	55	-	19
Other	1	-	-	0	-	0
TOTAL			102			35

TABLE VI-28

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR
Florida Site W

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.31	1.07	64	5.47	.90	30
Father's Education	5.36	1.36	52	5.82	1.22	28
Income	9.75	7.60	65	13.48	7.54	29
Occupation	15.28	2.74	65	15.70	2.02	30
Family Structure						
Mother and Father Present	75%	-	50	83%	-	25
Mother, not Father, Present	18	-	12	7	-	2
Other	7	-	5	10	-	3
TOTAL			67			30
Entry Level						
WRAT	29.11	12.93	66	24.77	11.26	33
PPVT	50.77	11.81	66	50.19	7.33	33
Preschool Experience						
Attended Headstart or other Preschool	36%	-	25	0%	-	0
TOTAL			69			31
Ethnicity						
Black	13%	-	8	0%	-	0
White	74	-	49	97	-	32
Spanish	9	-	6	0	-	0
Other	4	-	3	3	-	1
TOTAL			66			33

TABLE VI-29

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

Florida Site NC

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.58	1.01	43	5.61	.99	36
Father's Education	5.97	1.12	35	5.65	1.32	35
Income	15.64	6.51	42	15.77	6.78	35
Occupation	14.95	2.34	41	15.31	2.93	36
Family Structure						
Mother and Father Present	82%	-	35	91%	-	33
Mother, not Father, Present	16	-	7	6	-	2
Other	2	-	1	3	-	1
TOTAL			43			36
Entry Level						
WRAT	25.34	10.30	41	30.14	13.22	35
PPVT	47.49	7.90	41	50.77	10.78	35
Preschool Experience						
Attended Headstart or other Preschool	100%	-	17	missing		
TOTAL			17	missing		
Ethnicity						
Black	2%	-	1	0%	-	0
White	96	-	39	100	-	35
Spanish	0	-	0	0	-	0
Other	2	-	1	0	-	0
TOTAL			41			35

higher. Almost all of the FT children attended preschool; we have no preschool data for NFT.

In the small Far Western town (W), both FT and NFT children are predominantly White. The families of FT children have substantially lower incomes than NFT families. While none of NFT children attended preschool, 36% of the FT children did.

There are no large differences between FT and NFT at any of the three sites on kindergarten entry level tests. The FT children score somewhat higher than the NFT children in cases where the two groups are not roughly equal. Perhaps this is a result of the FT children's preschool experience.

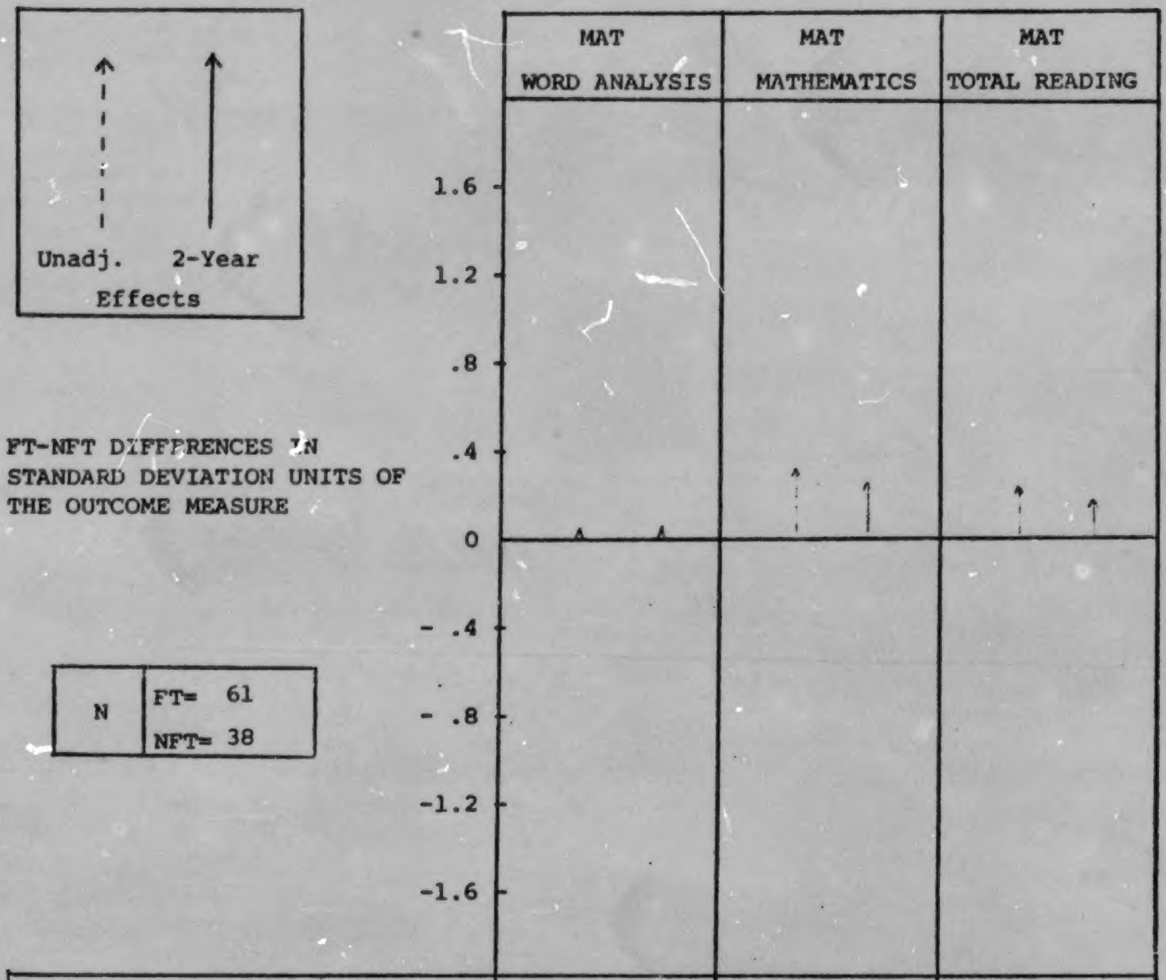
In the second type of site (NC, Table VI-29), both FT and NFT groups are nearly all White and essentially identical on measures of parental education, family income and family structure. In this small North Central community, both groups appear to come from intact homes, with somewhat higher SES than the FT families in the other Florida sites. However, the FT children's kindergarten entry level test scores are lower than those of the NFT children.

6.7.3 By Site 2-Year Effects

The FT/NFT differences for Florida, shown in Figures VI-34 through VI-39, show either no difference between FT and NFT or small FT-favoring differences in the three sites where the FT and NFT group are matched on entry scores. However, the results for NC, the site more poorly matched on entry scores, show either no difference or small NFT-favoring differences.

The pattern of effects and site characteristics for Florida is consistent with the interpretation that Florida's FT children are in general holding their ground, sometimes moving ahead and sometimes falling behind their more advantaged peers. Although the impact of preschool on pretest scores has not been assessed in this analysis, the Florida findings add evidence to the argument that preschool increases pretest scores, thereby leaving a bias in our analysis against those groups with a high percentage attending preschool.

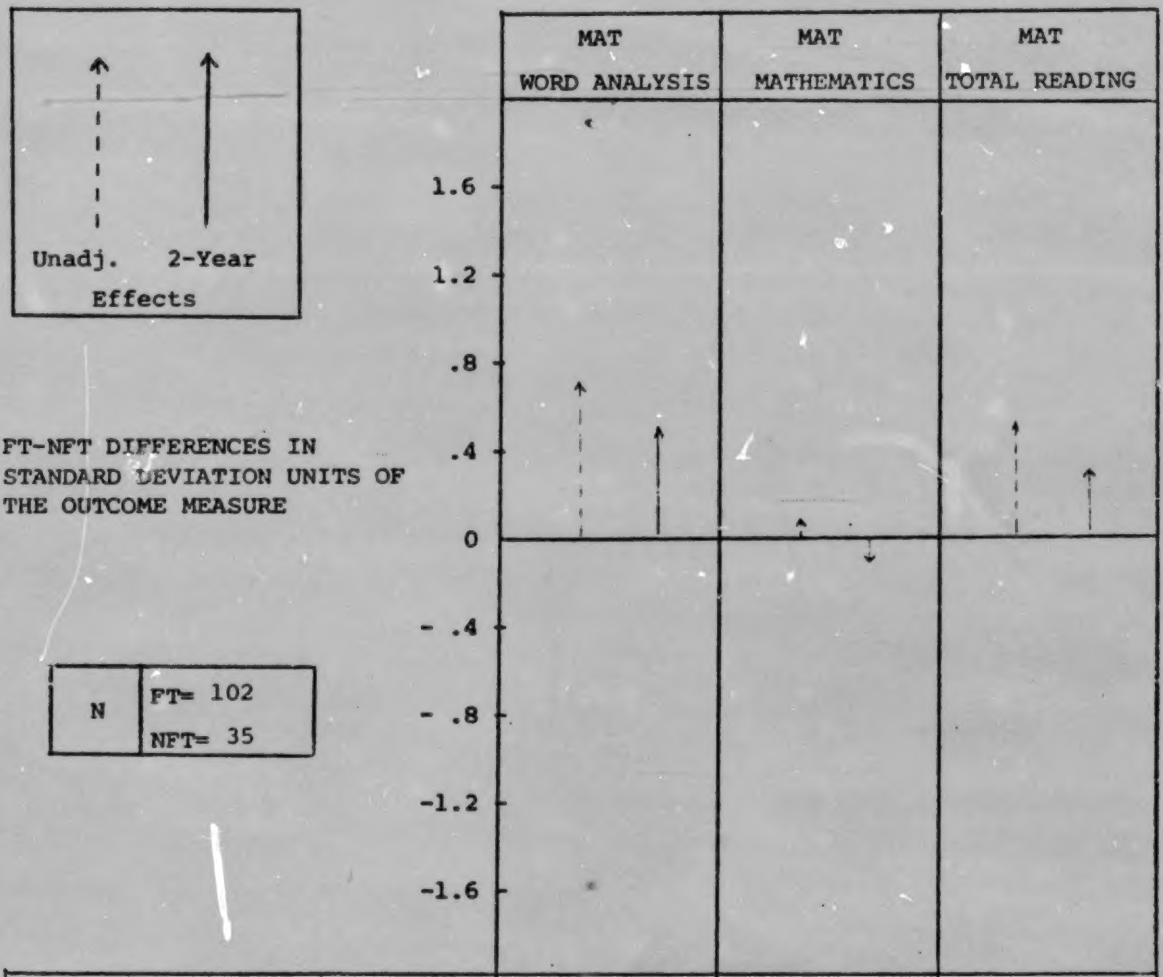
FIGURE VI-36 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FLORIDA SITE S1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.011	3.407	3.214
	2-Year	.002	2.750	2.362
S.E. = Standard Error of B	Unadjusted	.054	2.258	3.085
	2-Year	.050	1.990	2.848
F = Significance Statistic	Unadjusted	.042	2.277	1.086
	2-Year	.002	1.910	.687
SD=Raw Score Standard Deviation		.258	10.999	14.936
B/SD = Effect in SD Units (See Graph)	Unadjusted	.043	.310	.215
	2-Year	.008	.250	.158

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

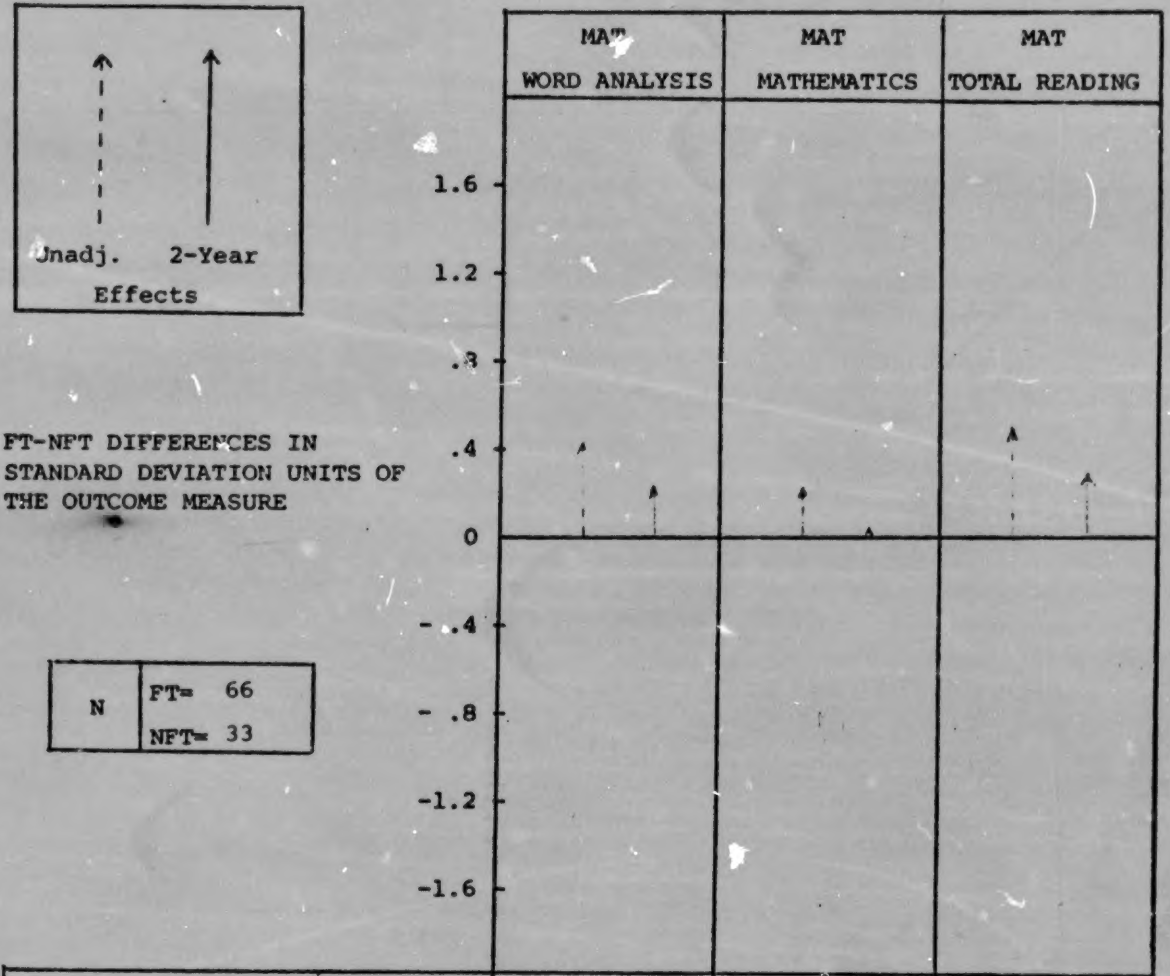
FIGURE VI-37 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FLORIDA SITE S2



B = Sponsor Effect	Unadjusted	.167	.686	6.765
(FT-NFT Differences)	2-Year	.120	- 1.162	4.025
S.E. = Standard	Unadjusted	.044	2.163	2.561
Error of B	2-Year	.038	2.030	2.248
F = Significance	Unadjusted	14.410*	.100	6.980*
Statistic	2-Year	9.973*	.327	3.204
SD=Raw Score Standard Deviation		.235	11.007	13.358
B/SD = Effect in SD	Unadjusted	.711	.062	.506
Units (See Graph)	2-Year	.511	- .106	.301

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

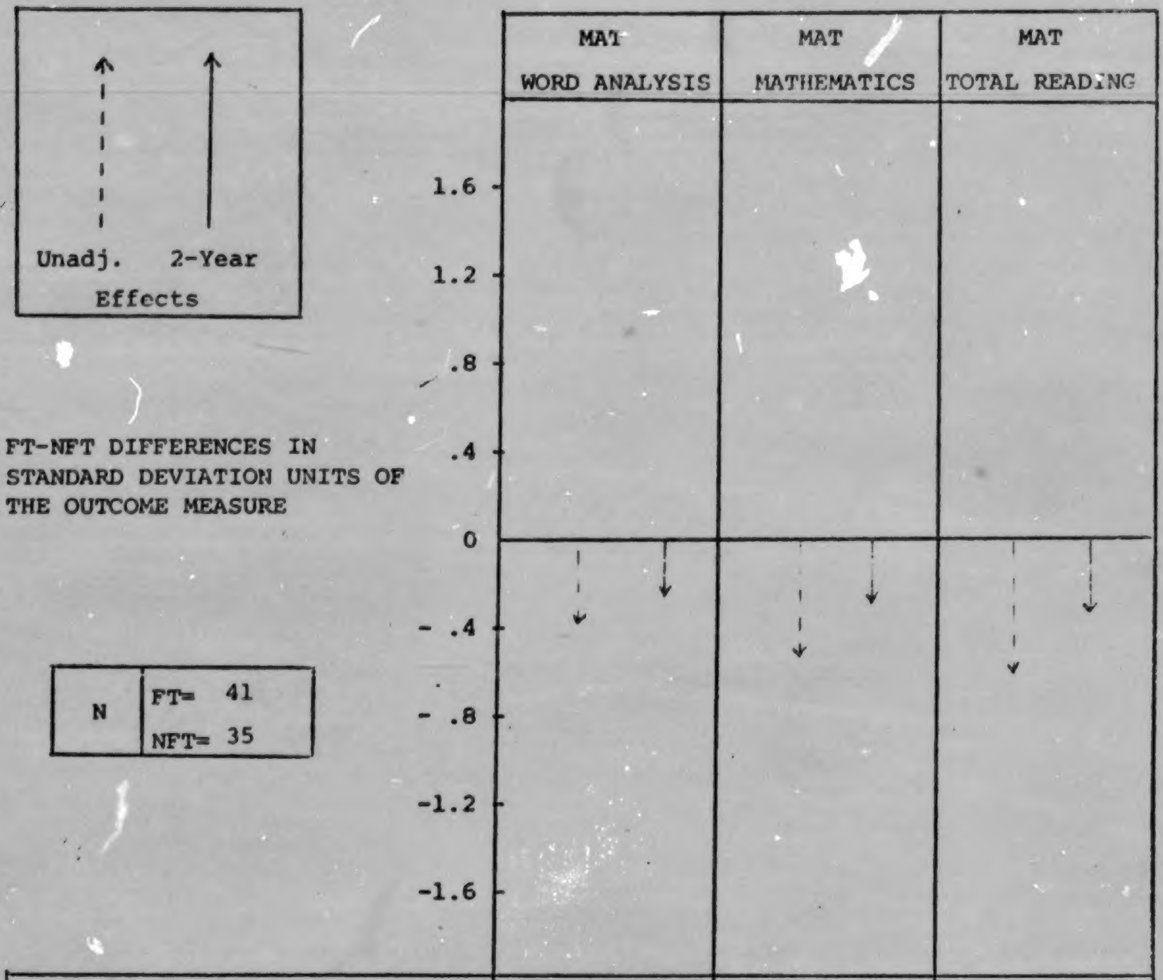
FIGURE VI-38 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FLORIDA SITE W



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect	Unadjusted	.107	2.470	8.297
(FT-NFT Differences)	2-Year	.058	.123	4.763
S.E. = Standard	Unadjusted	.055	2.395	3.584
Error of B	2-Year	.047	2.000	2.997
F = Significance	Unadjusted	3.787	1.063	5.359*
Statistic	2-Year	1.523	.004	2.525
SD=Raw Score Standard Deviation		.255	10.975	16.791
B/SD = Effect in SD	Unadjusted	.420	.225	.494
Units (See Graph)	2-Year	.228	.011	.284

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-39 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
FLORIDA SITE NC



N	
	FT= 41
	NFT= 35

		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .095	- 7.196	-10.556
	2-Year	- .040	- 3.804	- 5.771
S.E. = Standard Error of B	Unadjusted	.057	2.961	3.734
	2-Year	.049	2.222	2.500
F = Significance Statistic	Unadjusted	2.779	5.905*	7.992*
	2-Year	.666	2.931	5.327
SD=Raw Score Standard Deviation		.252	13.280	16.964
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .377	- .542	- .622
	2-Year	- .159	- .286	- .340

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

6.8 THE EDUCATIONAL DEVELOPMENT CENTER (EDC)

The EDC program is based on the belief that learning is facilitated by a child's active participation in the learning process. In the flexible, open classroom environment, children are encouraged to initiate activities, pursue their interests, and generally assume responsibility for their own learning. The basic objective of the program is to provide the optimal environment for children's growth in academic and problem solving skills, self-expression, and self-direction.

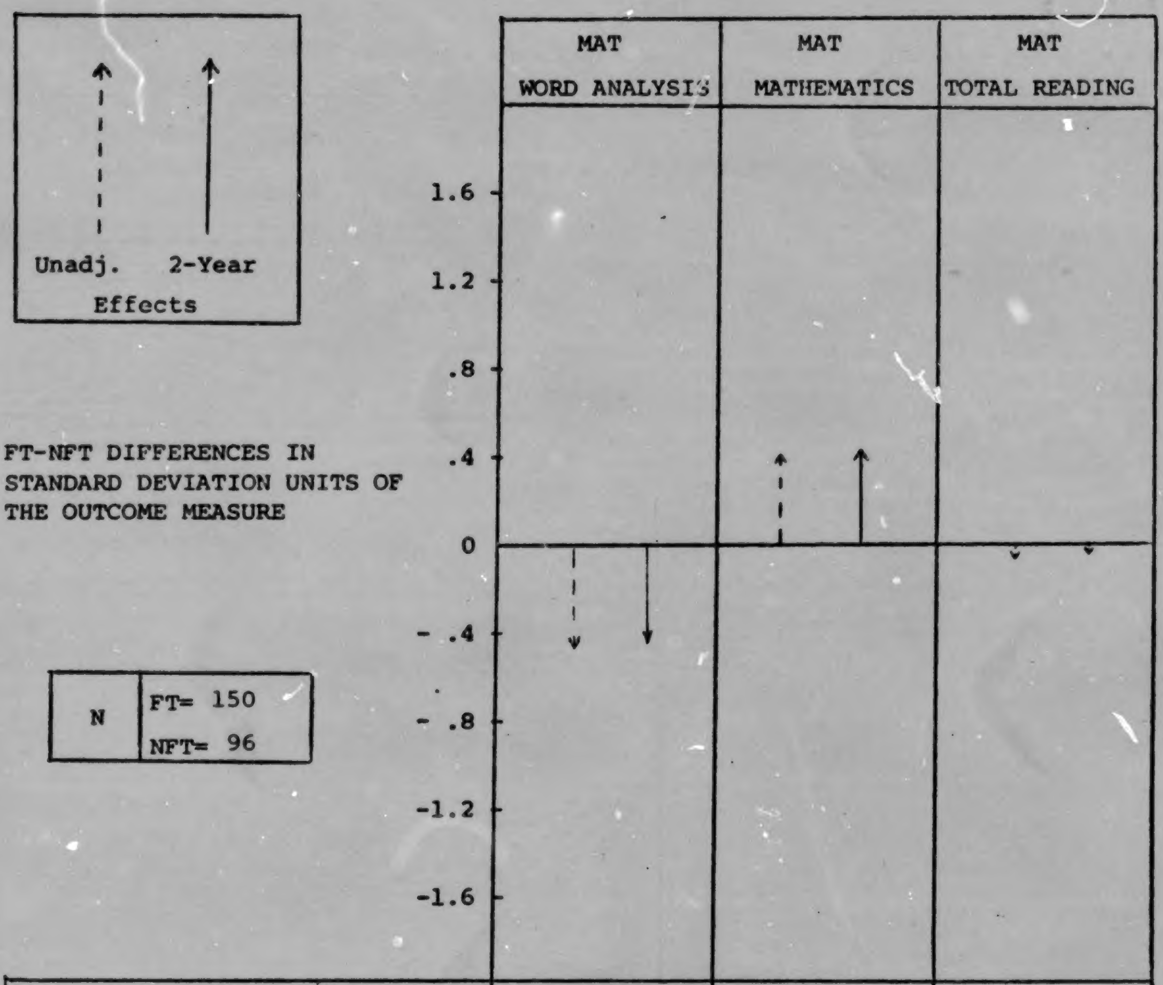
6.8.1 Overall 2-Year Effects

Figure VI-40 shows that, at the end of first grade, there is a strong FT effect on Mathematics, no difference in Total Reading effects between FT and NFT, and a strong NFT effect in Word Analysis. The FT and NFT groups are generally well matched at the overall Sponsor level; this is not true when the sites are considered separately. These effects can therefore be attributed to the EDC program with some degree of confidence. The most outstanding difference between the groups is in the number of FT children who attended preschool prior to entering kindergarten. The FT children had a much higher rate of such experiences than the NFT children, and this may confound the EDC effects during FT. Nevertheless, the FT effect in Mathematics is very strong and reasonably related to the nature of the program. At the same time, the negative effects in Word Analysis are consistent with this model's avoidance of traditional academic materials and topics. Although it is not the purpose of this chapter to contrast the one and two-year effects (the interested reader is referred to Monograph I, Vol. IIB of this report for a discussion of this topic), it is worth noting that these strong mathematics effects are not found after a year of kindergarten. There is likely to be a unique effect of the EDC model in first grade. There is also some variability in these findings when they are examined at each site. We turn now to the individual site descriptions.

6.8.2 Site Descriptions

This evaluation of the EDC program examines data from three sites, all large cities in the Northeast and Middle Atlantic areas. The FT groups in general have lower kindergarten entry scores and/or lower SES than their NFT counterparts. The family and pupil characteristics are shown in Tables VI-30 through VI-32.

FIGURE VI-40 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
EDC



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .125	5.371	- 1.086
	2-Year	- .119	5.630	- .753
S.E. = Standard Error of B	Unadjusted	.034	1.704	2.113
	2-Year	.025	1.572	1.668
F = Significance Statistic	Unadjusted	13.814*	9.935	.264
	2-Year	22.173*	12.826	.204
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .468	.410	- .062
	2-Year	- .447	.430	- .043

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

TABLE VI-30

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.96	1.03	46	5.94	.80	35
Father's Education	5.81	1.22	42	6.09	.63	33
Income	13.67	6.58	45	14.71	6.18	34
Occupation	14.27	2.20	44	15.14	1.59	35
Family Structure						
Mother and Father Present	85%	-	39	91%	-	33
Mother, not Father, Present	9	-	4	6	-	2
Other	6	-	3	3	-	1
TOTAL			46			36
Entry Level						
WRAT	33.23	12.54	69	38.38	11.58	39
PPVT	50.56	7.69	69	53.31	6.50	39
Preschool Experience						
Attended Headstart or other Preschool	60%	-	35	10%	-	4
TOTAL			58			40
Ethnicity						
Black	0%	-	0	0%	-	0
White	100	-	69	100	-	39
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			69			39

TABLE IV-31

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.14	.88	50	4.77	.84	31
Father's Education	4.91	1.12	23	5.07	1.10	15
Income	7.50	5.50	50	9.91	6.06	33
Occupation	15.94	1.05	49	15.89	1.50	28
Family Structure						
Mother and Father Present	45%	-	23	42%	-	14
Mother, not Father, Present	55	-	28	49	-	16
Other	0	-	0	9	-	3
TOTAL			51			33
Entry Level						
WRAT	24.13	11.46	45	24.34	10.32	29
PPVT	34.60	14.72	45	37.21	14.58	29
Preschool Experience						
Attended Headstart or other Preschool	87%	-	46	21%	-	7
TOTAL			53			33
Ethnicity						
Black	98%	-	44	82%	-	24
White	0	-	0	3	-	1
Spanish	2	-	1	15	-	4
Other	0	-	0	0	-	0
TOTAL			45			29

TABLE VI-32

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.18	.66	22	5.23	.88	31
Father's Education	missing			5.54	.69	11
Income	9.77	5.57	22	9.85	7.25	33
Occupation	16.00	1.94	19	16.17	1.64	30
Family Structure						
Mother and Father Present	46%	-	10	36%	-	12
Mother, not Father, Present	50	-	11	61	-	20
Other	4	-	1	3	-	1
TOTAL			22			33
Entry Level						
WRAT	19.14	6.69	24	26.65	12.14	16
PPVT	35.36	8.14	24	35.77	9.28	16
Preschool Experience						
Attended Headstart or other Preschool	100%	-	22	0%	-	0
TOTAL	100%	-	22			33
Ethnicity						
Black	100%	-	24	97%	-	15
White	0	-	0	3	-	1
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			24			16

At site NE1 (Table VI-30), a small Northeastern city, both FT and NFT children are all White, but the FT children have slightly lower SES. The FT children are behind the NFT children on kindergarten entry tests, and a considerably higher percentage of FT children attended preschool (FT = 60%; NFT = 10%). Substantially more Title 1 money is expended in the NFT schools, suggesting that a compensatory curriculum is operating in these schools.

At site NE2 (Table VI-31), a large Northeastern city, both FT and NFT groups are predominantly Black (FT = 98%; NFT = 82%, 15% Spanish-surnamed). FT and NFT are similar on measures of SES, except that FT family income is lower. More FT children than NFT children attended preschool (FT = 87%; NFT = 21%); however, the two groups show no differences on kindergarten entry level tests. One surprising difference between the groups is that the NFT first grade classes are considerably smaller, 16 children compared with 26 children in the FT classes. This contrast threatens the validity of any comparison of these groups since class size was not used as an adjusting factor in our analyses.

The remaining EDC site (S)[1] (Table VI-32), is a large Middle Atlantic city. Both FT and NFT groups are almost entirely Black and very similar on all SES indicators available. However, the FT group is a full standard deviation behind the NFT group on the kindergarten entry WRAT. In addition, 100 percent of the FT children and none of the NFT children had preschool experience. In this site, as in the others, there is a very weak match of FT and NFT groups.

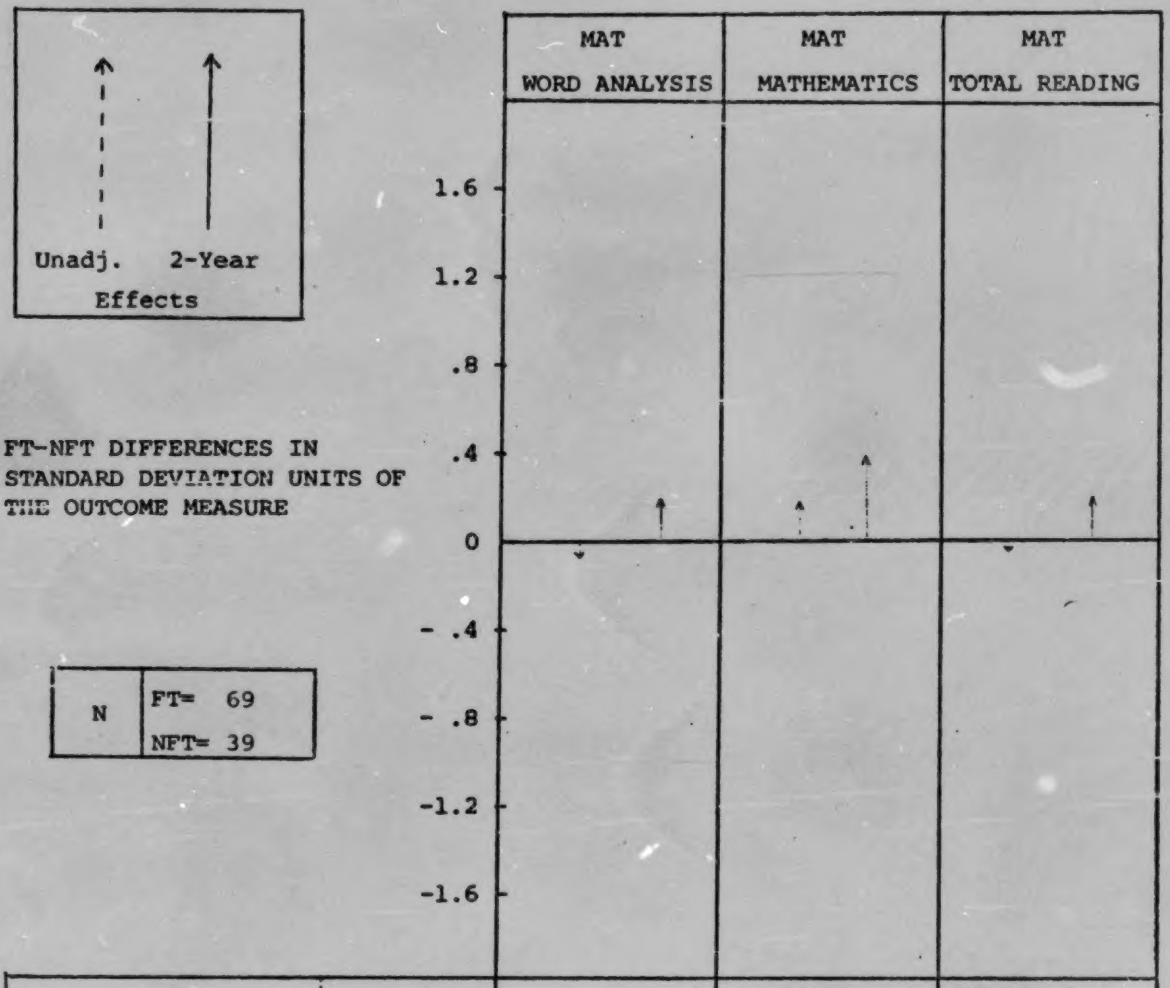
6.8.3 By Site 2-Year Effects

Despite three rather different groups of children at these three sites, there is a consistent and strong Mathematics effect in all three. The remaining effects appear to be dependent upon the character of the children at the sites. For site NE1 (Figure VI-41) in which the FT children start out at much lower levels of performance and have much lower SES than the NFT children, the Mathematics effect is the only one in evidence; FT and NFT are essentially equal on the other two outcomes. Site NE2 (Figure VI-42) has a better match,

---

[1] Kindergarten absence data were not available for this site; therefore, none of these children were included in the overall aggregated analysis either last year or this year. They do appear in the present site-by-site analysis. It should be noted also that this was a Head Start Planned Variation site, a factor which may contribute to its uniqueness.

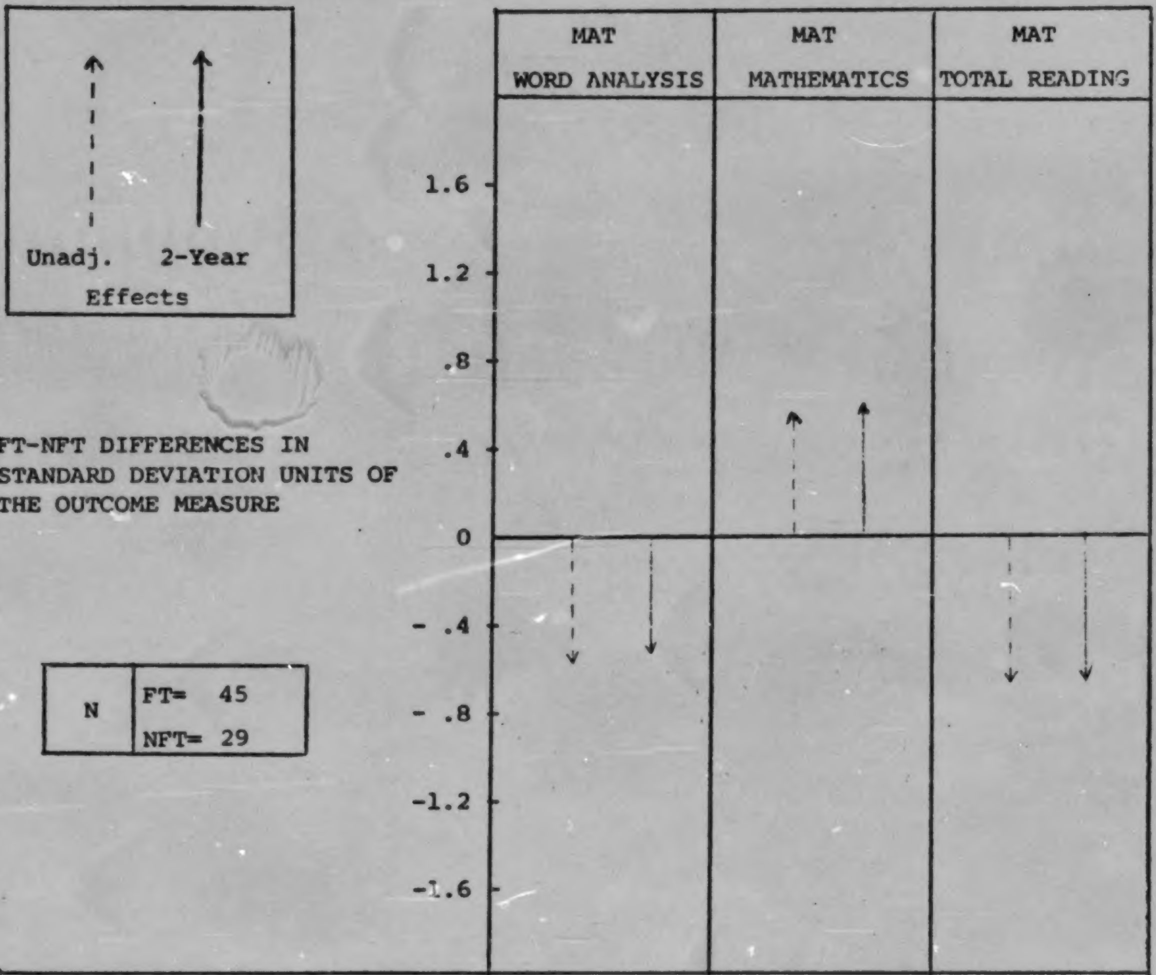
FIGURE VI-41 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
EDC SITE NE1



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .019	1.656	- .949
	2-Year	.049	3.925	3.248
S.E. = Standard Error of B	Unadjusted	.050	2.124	3.353
	2-Year	.043	1.898	2.885
F = Significance Statistic	Unadjusted	.144	.608	.080
	2-Year	1.300	4.277*	1.268
SD=Raw Score Standard Deviation		.247	10.581	16.663
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .077	.156	- .057
	2-Year	.198	.371	.195

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

FIGURE VI-42 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
EDC SITE NE2

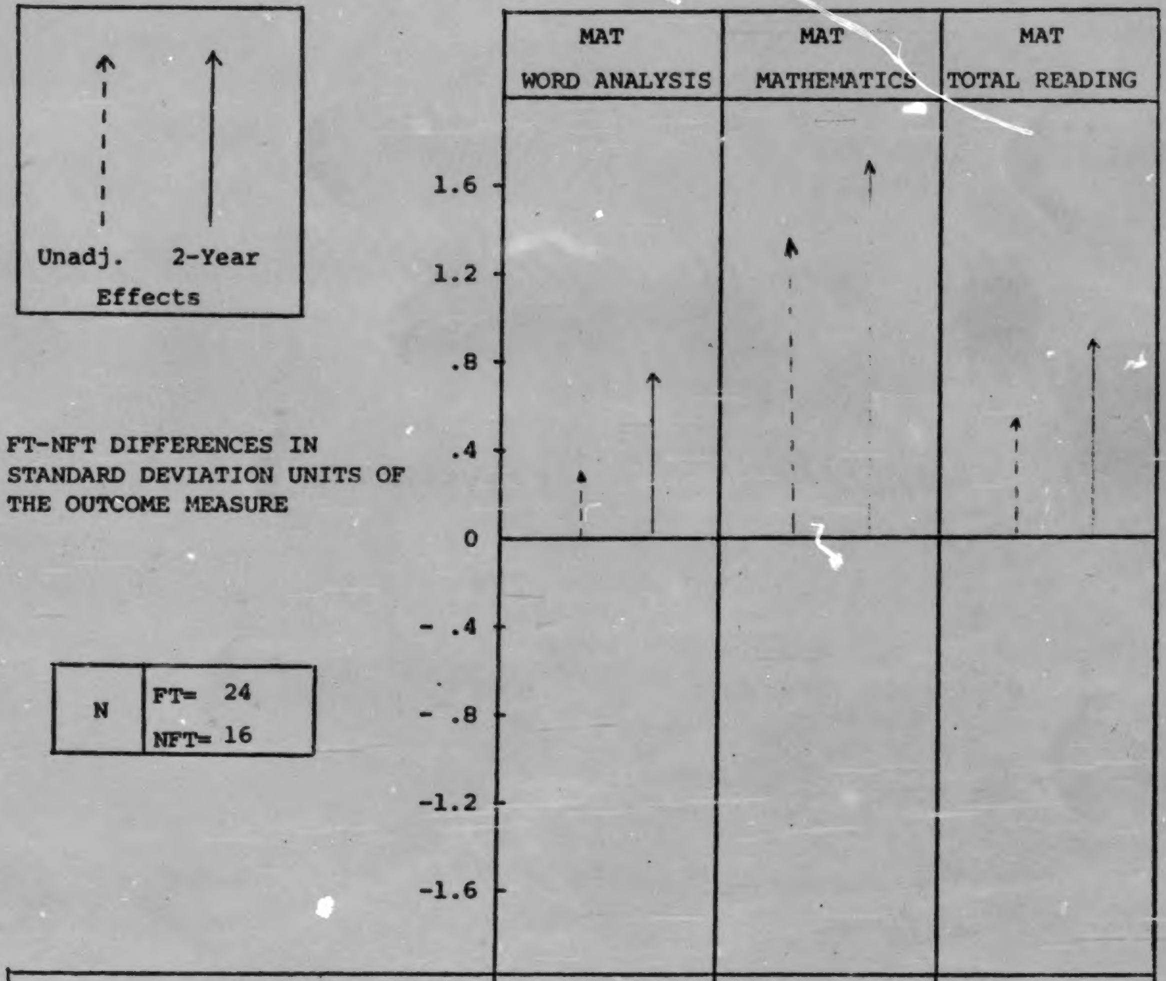


		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .140	5.852	- 9.263
	2-Year	- .128	6.160	- 9.379
S.E. = Standard Error of B	Unadjusted	.054	2.450	3.079
	2-Year	.052	2.251	2.718
F = Significance Statistic	Unadjusted	6.724*	5.707*	9.048*
	2-Year	6.061*	7.491*	11.909*
SD=Raw Score Standard Deviation		.237	10.616	13.624
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .591	.551	- .680
	2-Year	- .540	.580	- .688

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's educatio
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

FIGURE VI-43 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE

EDC SITE S



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.069	18.962	8.198
	2-Year	.170	24.139	13.755
S.E. = Standard Error of B	Unadjusted	.077	3.715	5.108
	2-Year	.069	3.137	4.528
F = Significance Statistic	Unadjusted	.803	35.665*	2.576
	2-Year	6.071*	59.198*	9.229*
SD=Raw Score Standard Deviation		.227	14.135	15.324
B/SD = Effect in SD Units (See Graph)	Unadjusted	.304	1.342	.535
	2-Year	.749	1.708	.898

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

but the site has some puzzling characteristics; the NFT class sizes are only about half the size of the FT classes. Here the Mathematics effects are paired with strong negative effects on Total Reading and Word Analysis. In site S (Figure VI-43), a well matched site, but one in which the EDC Follow Through children all attended an EDC Head Start program before entering kindergarten, all three outcome measures show very strong positive effects.

There is clear evidence of strong effects across sites for this Sponsor, a finding not to be seen at the end of kindergarten. EDC, along with two or three other Sponsors, appears to be offering one effective means toward increased academic performance.

6.9 UNIVERSITY OF PITTSBURGH, PRIMARY EDUCATION PROJECT

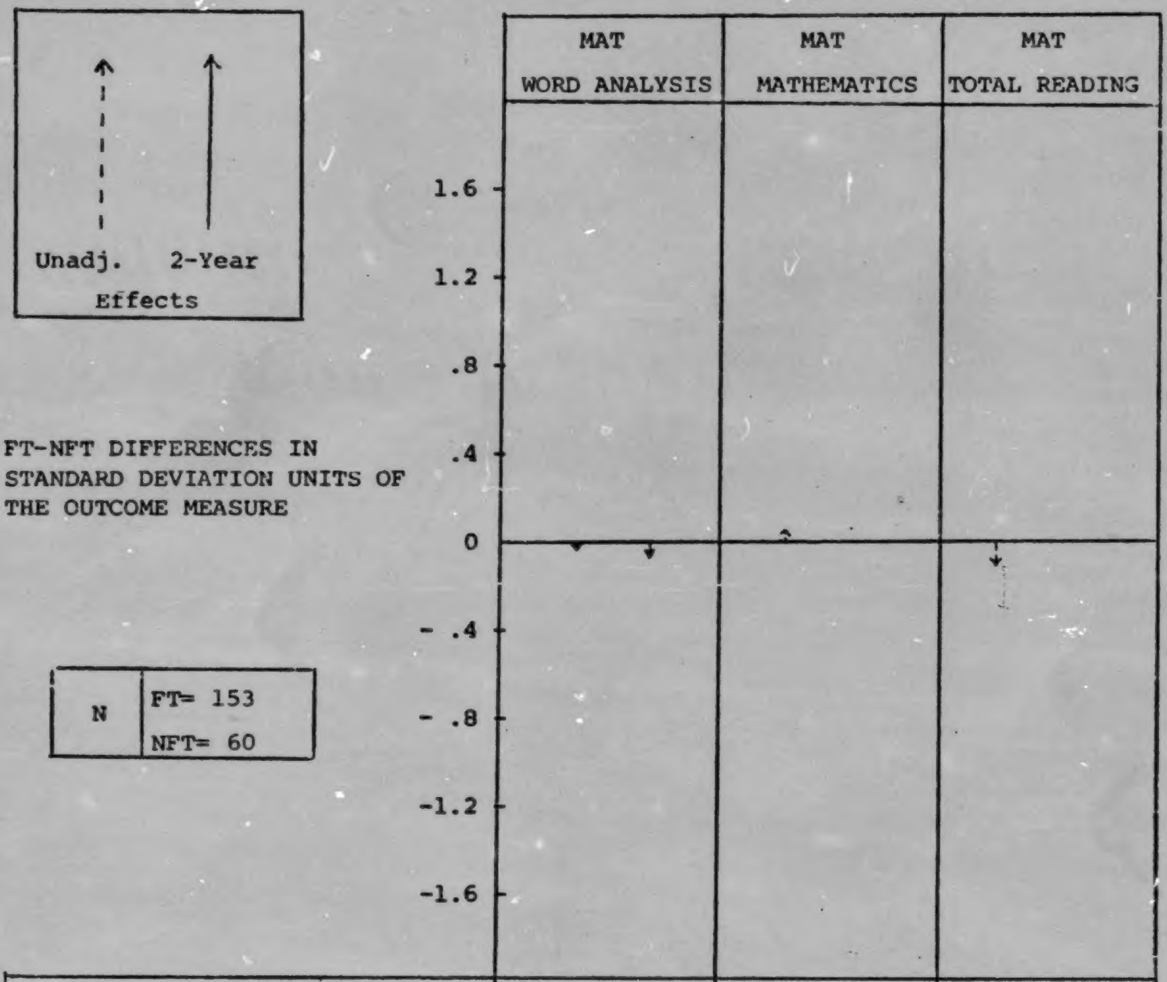
The Primary Education Project utilizes a number of interrelated curriculum components which are carefully structured and sequenced to provide for optimally efficient learning. Three general classes of skills are included in these curriculum components which are designed to form the foundation of all higher level functioning: (1) orienting and attending skills, (2) perceptual motor skills, and (3) conceptual-linguistic skills. The latter include classification, reasoning, memory, language, and mathematics concepts. The curriculum is highly individualized in order to allow the child to progress at his own pace. The teacher serves as a facilitator and resource person as the child moves through each component.

6.9.1 Overall 2-Year Effects

The FT/NFT differences for Pittsburgh are shown in Figure VI-44. Only the effects for Word Analysis are shown; there is no difference between FT and NFT. Only unadjusted differences for Total Reading and Mathematics are shown, since the relationship between these outcomes and the entry WRAT were heterogeneous across FT and NFT. The relationship in the two groups for Total Reading, shown in Figure VI-45, indicates no FT/NFT difference on either entry WRAT or Total Reading. However, the difference between the regressions indicates that NFT children with lower scores at entry gained somewhat more than FT children with comparable scores. The Mathematics relationships, shown in Figure VI-46, also show no difference on pretest or posttest means and indicate that the lower entry level NFT children scored slightly higher than lower entry level FT children, and also that higher entry level FT children scored slightly higher than NFT children with comparable entry scores. These different relationships suggest that the Pittsburgh program is differentially effective depending on the initial achievement level of the children, working somewhat better with higher entry level children, at least in mathematics. The Pittsburgh group is among the least disadvantaged in the entire analysis, having relatively high values on both SES indicators and entry tests.

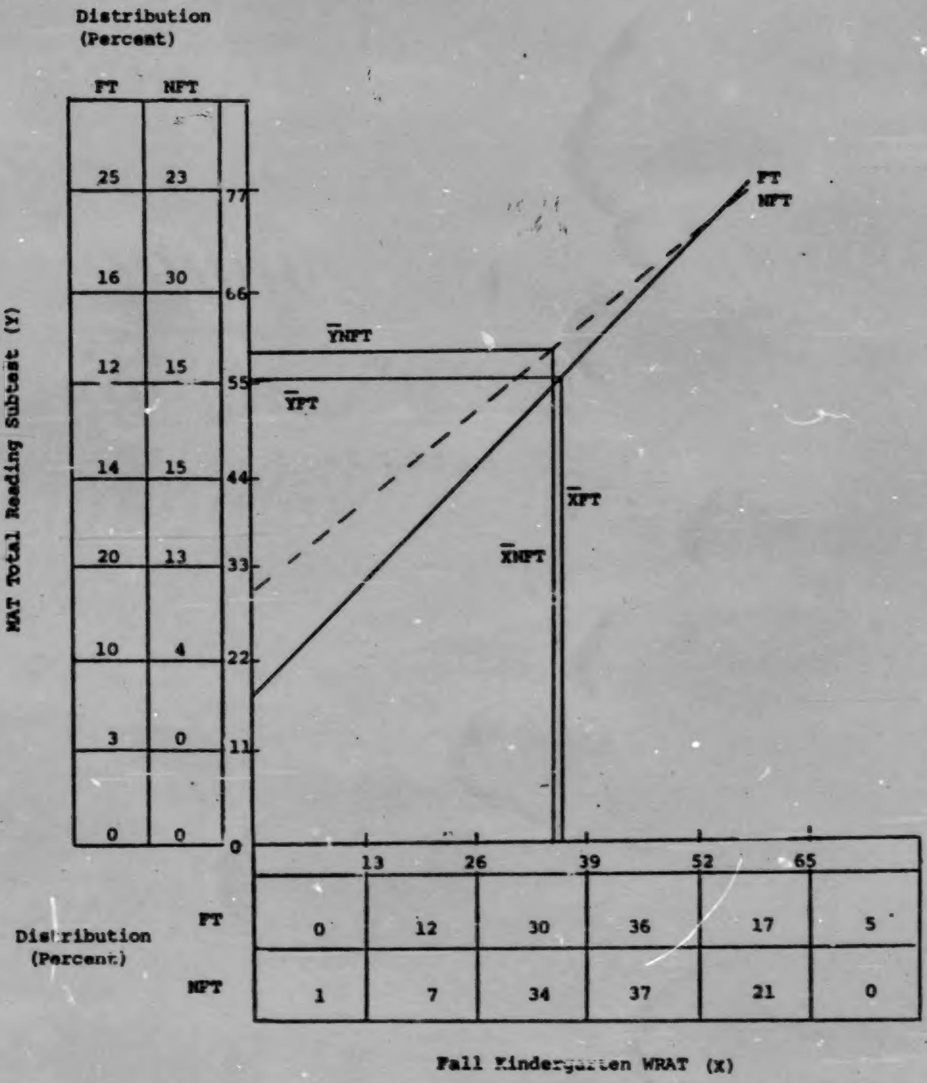
The findings suggest that the Pittsburgh curriculum does not add substantially to the academic achievement of children participating in

FIGURE VI-44 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
PITTSBURGH



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .007	.554	- 3.612
	2-Year	- .017		
S.E. = Standard Error of B	Unadjusted	.038	2.070	2.766
	2-Year	.029		
F = Significance Statistic	Unadjusted	.034	.072	1.705
	2-Year	.355		
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .026	.042	- .206
	2-Year	- .065		

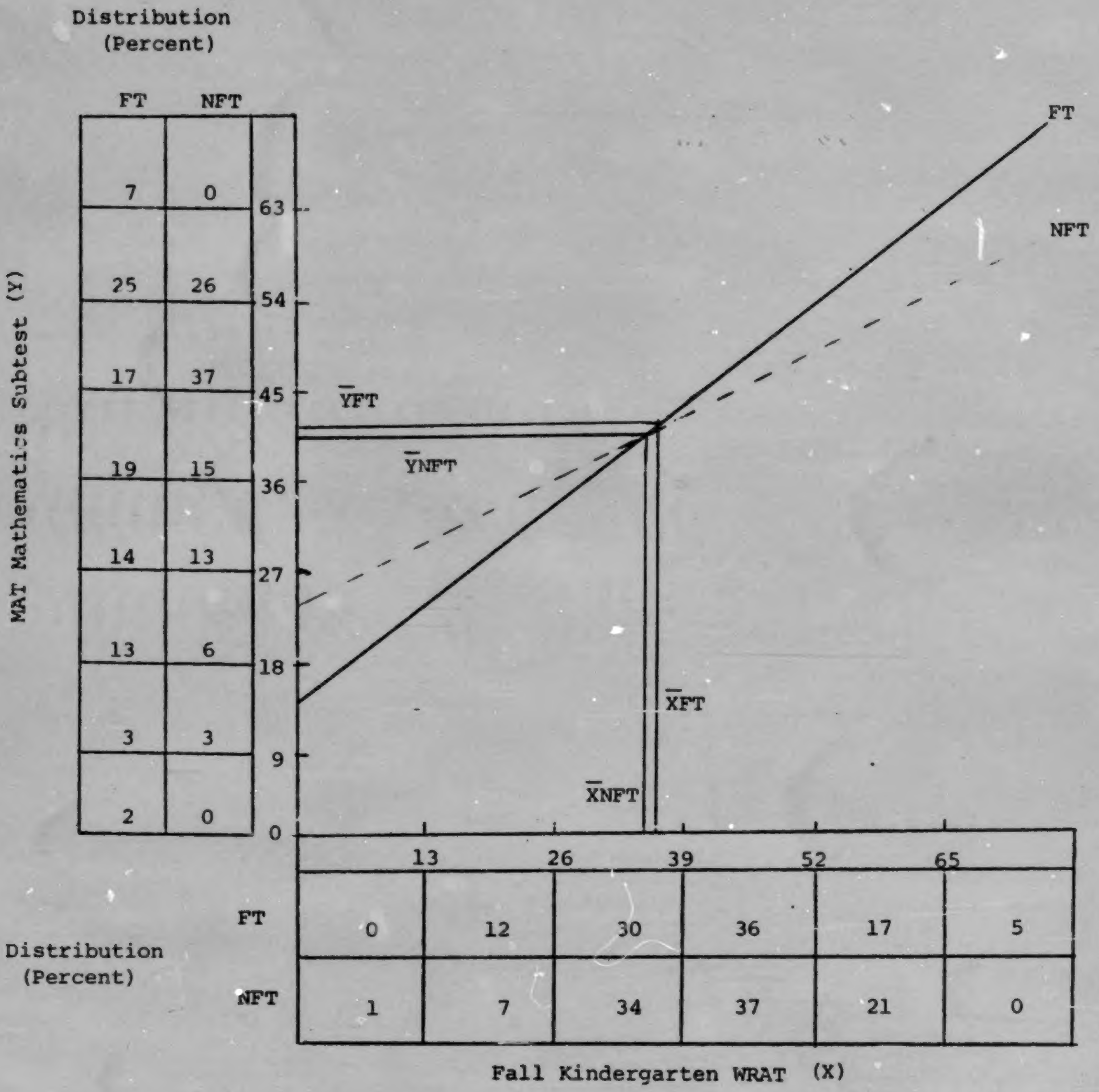
NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05



	FT	NFT
X	36.25	35.42
Y	55.25	58.87
b	1.08	.84
SE (b)	.09	.14
n	153	60

Figure VI-45

Heterogeneous Regressions:
Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT
Sponsor: Pittsburgh
Outcome: MAT Total Reading Subtest



FT	NFT
7	0
25	26
17	37
19	15
14	13
13	6
3	3
2	0

FT	0	12	30	36	17	5
NFT	1	7	34	37	21	0

	FT	NFT
$\bar{X}$	36.25	35.42
$\bar{Y}$	41.47	40.92
b	.76	.60
SE (b)	.07	.11
n	153	60

Figure VI-46

Heterogeneous Regressions:
Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT

Sponsor: Pittsburgh
Outcome: MAT Mathematics Subtest

the program. The program appears to be just as good but no better than curriculum experienced by the NFT children. Whether the Pittsburgh program is in part less effective with lower achieving children, as the Mathematics and Total Reading results suggest , cannot be conclusively determined from the overall analysis. However, the site analysis will illuminate this issue.

6.9.2 Site Description

Only two out of eight Pittsburgh sites met the criteria for inclusion in these analyses. These were a rural community in the North Central region of the country (NC, Table VI-33) and a small Northeastern city (NE, Table VI-34). The assessment of the effectiveness of a program based on only two sites has inherent weaknesses. Furthermore, the two sites in which there are sufficient numbers of tested children differ substantially from one another on a variety of pupil characteristics as well as on the nature of the FT/NFT match within site. These differences prevent examination of Sponsor effects independent of specific site or child variables. Thus, the findings for this Sponsor must be interpreted with caution.

The FT group in the North Central community (NC) is one of the highest in the analytic sample with respect to family income, level of parental education, occupational status, and entry scores. Most of the children are White, come from intact families, and have attended preschool programs. The NFT group in this site is very similar to FT on both SES and other characteristics.

The FT group in the Northeastern site (NE) is substantially lower than the NFT group on SES, has a lower percentage of intact families, and has lower pretest scores. The majority of the FT group also attended preschool (79%) whereas none of the NFT group did.

The FT group in this site is about average in terms of its status in the overall analysis, much lower than site NC.

6.9.3 By Site 2-Year Effects

The FT/NFT differences on end of first grade measures for Pittsburgh's sites, shown in Figures VI-47 and VI-48, indicate no difference between FT and NFT in site NC where the groups are well matched but have higher entry scores. However, site NE shows substantial NFT-favoring

TABLE VI-33

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

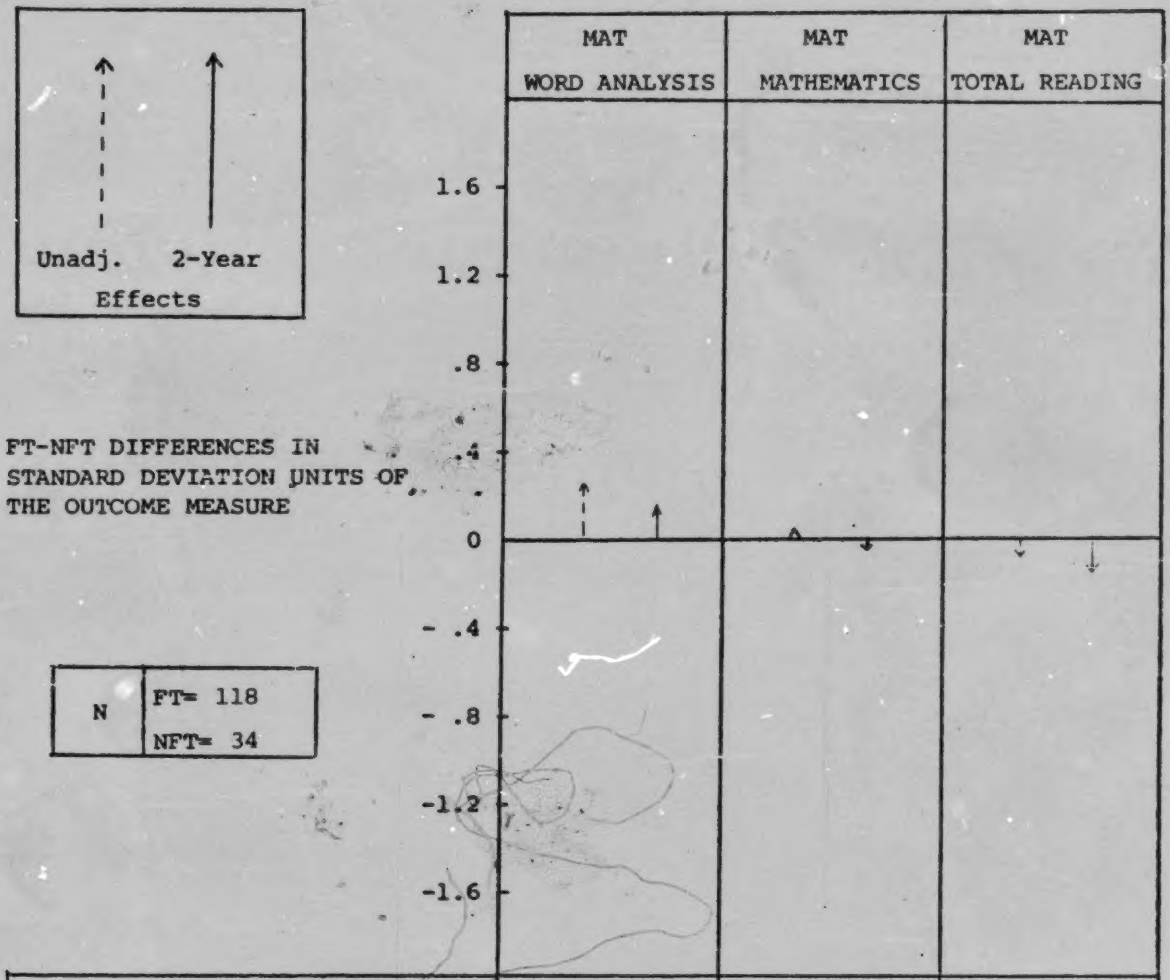
	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	6.08	.90	111	5.89	.92	36
Father's Education	6.03	1.18	107	5.83	1.23	36
Income	15.62	6.41	105	14.11	6.68	36
Occupation	13.68	2.55	111	12.89	2.67	36
Family Structure						
Mother and Father Present	93%	-	104	97%	-	35
Mother, not Father, Present	4	-	5	0	-	0
Other	3	-	3	3	-	1
TOTAL			112			36
Entry Level						
WRAT	39.19	11.45	118	37.76	10.75	34
PPVT	56.70	6.01	118	56.41	5.89	34
Preschool Experience						
Attended Headstart or other Preschool	97%	-	31	100%	-	15
TOTAL			32			15
Ethnicity						
Black	0%	-	0	0%	-	0
White	98	-	116	97	-	33
Spanish	0	-	0	0	-	0
Other	2	-	2	3	-	1
TOTAL			118			34

TABLE VI-34

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	5.42	.64	60	6.00	.95	30
Father's Education	5.35	.77	40	5.77	1.11	26
Income	7.89	5.70	62	14.50	6.17	30
Occupation	16.28	1.29	60	15.16	2.00	31
Family Structure						
Mother and Father Present	51%	-	32	84%	-	26
Mother, not Father, Present	31	-	20	13	-	4
Other	18	-	11	3	-	1
TOTAL			63			31
Entry Level						
WRAT	28.68	10.86	66	32.37	13.52	30
PPVT	47.06	8.22	66	51.47	11.46	30
Preschool Experience						
Attended Headstart or other Preschool	79%	-	61	0%	-	0
TOTAL			77			32
Ethnicity						
Black	0%	-	0	0%	-	0
White	100	-	66	100	-	30
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			66			30

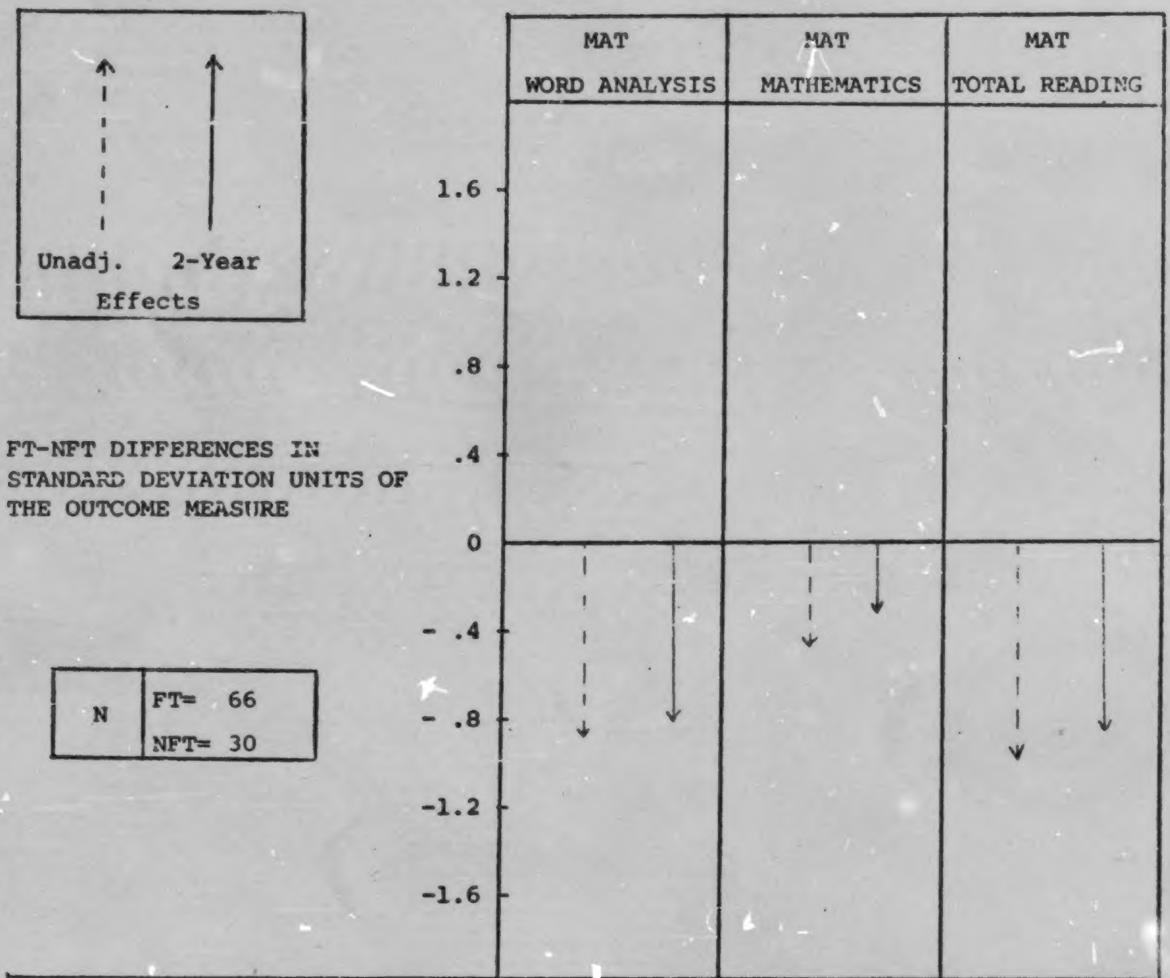
FIGURE VI-47 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
PITTSBURGH SITE NC



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.032	.114	- 1.008
	2-Year	.019	- .494	- 2.073
S.E. = Standard Error of B	Unadjusted	.027	1.792	2.495
	2-Year	.024	1.618	2.027
F = Significance Statistic	Unadjusted	1.404	.004	.163
	2-Year	.627	.093	1.046
SD=Raw Score Standard Deviation		.139	9.178	12.783
B/SD = Effect in SD Units (See Graph)	Unadjusted	.230	.012	- .079
	2-Year	.137	- .054	- .162

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

FIGURE VI-48 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
PITTSBURGH SITE NE



B = Sponsor Effect	Unadjusted	- .213	- 5.258	-16.200
(FT-NFT Differences)	2-Year	- .195	- 3.535	-14.004
S.E. = Standard	Unadjusted	.048	2.406	3.159
Error of B	2-Year	.042	1.844	2.645
F = Significance	Unadjusted	19.696*	4.774*	26.296*
Statistic	2-Year	21.558*	3.675	28.026*
SD=Raw Score Standard Deviation		.240	11.142	16.144
B/SD = Effect in SD	Unadjusted	- .888	- .472	- 1.004
Units (See Graph)	2-Year	- .812	- .317	- .867

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

differences on all outcomes. Here, the FT children are below NFT on all background indicators and scored substantially below NFT on entry tests. The FT children in site NE appear to be falling progressively further behind their more advantaged peers.

The use of caution in the interpretation of the Pittsburgh model results cannot be overemphasized. Only two sites are represented in these findings, one characterized by an unusually high FT/NFT group in terms of both SES and entering achievement, and the other by a severe FT/NFT mismatch. However, the differences between the effects in the sites correspond to the results found in the overall analysis. The higher initial achieving group, site NC, shows no particular advantage from participation in the Pittsburgh program whereas the lower achieving group, albeit average for this analysis, shows a loss from participation in the program. The Pittsburgh program does not appear to be meaningfully influencing the expected progress of the children in either site.

6.10 SOUTHWEST EDUCATION DEVELOPMENT LABORATORY (SEDL)

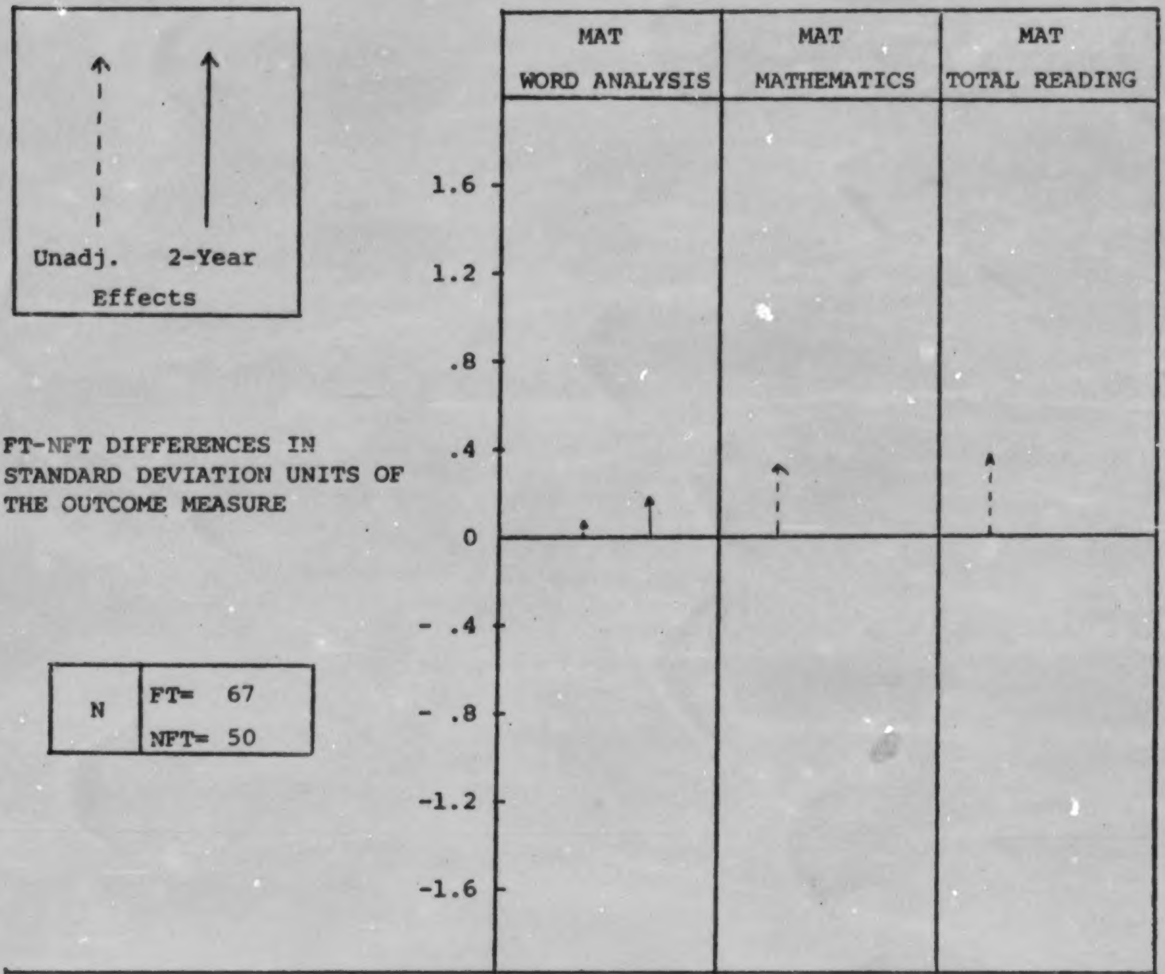
The Language Development (Bilingual and/or English-as-a-second language) Approach was originally designed as an instructional program for predominantly Spanish-speaking classrooms. The primary emphasis of the approach is on language development; language is seen as the tool for acquiring a variety of skills including non-linguistic skills. Building upon the child's native language and culture, the kindergarten program stresses the development of visual, auditory, and motor skills, as well as thinking, discovery, and English language structures.

6.10.1 Overall 2-Year Effects

Two of the three outcome measures (Total Reading and Mathematics) could not be examined with a covariance adjustment at the overall Sponsor level because of a heterogeneous relationship between the covariable and the outcome measures for the FT and NFT group. The third outcome measure, Word Analysis, showed no difference between the FT and NFT groups after two years in the program (see Figure VI-49).

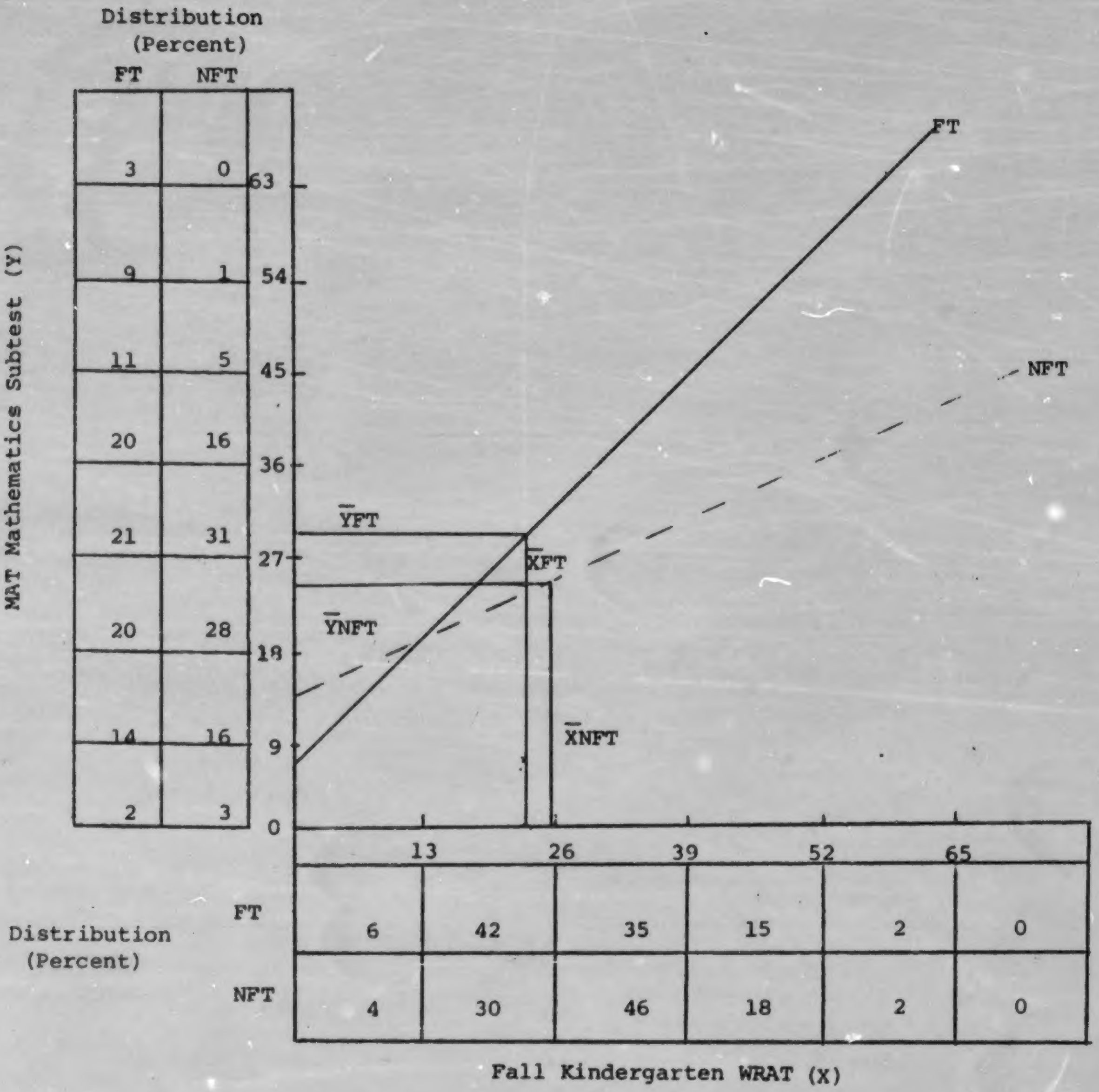
Further examination of these heterogeneous findings can be made by referring to Figure VI-50 and Figure VI-51. Here the unadjusted differences in Mathematics and Total Reading are plotted against the covariable (Fall Kindergarten WRAT). It is clear that the FT children at the higher end of the pretest distribution are performing at a much higher level in both mathematics and reading than the NFT children at the same high pretest level. The FT children who did less well on the pretest, on the other hand, are performing in reading and mathematics at the same level as the NFT children who did less well on the pretest. Thus, the heterogeneity reflects the SEDL program's apparent disproportional impact upon higher achieving children. This heterogeneity also disappears when the data are organized and analyzed at site level which suggests that sites may differ from each other on the distribution of the covariable, or they may differ from each other in the distribution of first grade Total Reading and Mathematics scores. In either case, the site level analyses, to which we now turn, reveal a good deal more than the overall Sponsor analysis.

FIGURE VI-49 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
SEDL



		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.017	4.339	6.539
	2-Year	.048		
S.E. = Standard Error of B	Unadjusted	.036	2.161	1.830
	2-Year	.031		
F = Significance Statistic	Unadjusted	.230	4.032	12.768*
	2-Year	2.353		
SD=Raw Score Standard Deviation		.267	13.092	17.533
B/SD = Effect in SD Units (See Graph)	Unadjusted	.065	.331	.373
	2-Year	.179		

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

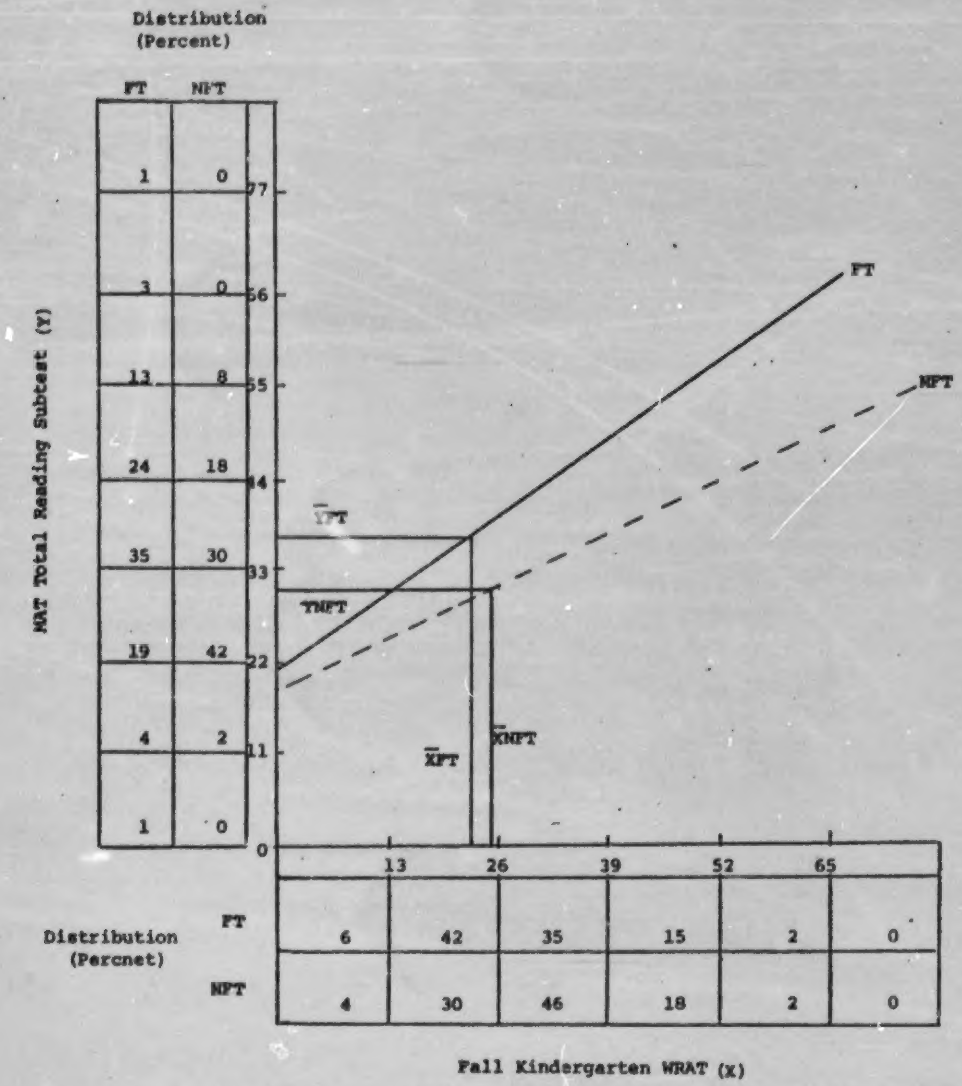


	FT	NFT
X̄	23.09	25.80
Ȳ	29.18	24.84
b	.87	.55
SE (b)	.12	.14
n	67	50

Figure VI-50

Heterogeneous Regressions:
Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT

Sponsor: SEDL
Outcome: MAT Mathematics Subtest



	FT	NFT
$\overline{X}$	23.09	25.80
$\overline{Y}$	37.18	30.64
b	.56	.29
SE (b)	.11	.14
n	67.	50

Figure VI-51

Heterogeneous Regressions:
Pre- and Post-Test Means, Distributions, and Regressions for FT and NFT
Sponsor: SEDL
Outcome: MAT Total Reading Subtest

6.10.2 Site Descriptions

This evaluation of the SEDL program examines data from three sites, one in a large Northeastern city (NE, Table VI-35), the other two in small towns in the West (W, Table VI-36) and South (S, Table VI-37). The FT and NFT children within these sites differ somewhat in demographic characteristics.

The characteristic shared by all of the SEDL sites is their substantial non-English speaking population. Two of them (NE and W) have a large Spanish-speaking population, while the remaining site (S) has a large number of people whose primary language is Cajun French. Aside from this basic similarity, the sites differ in several important respects.

At site NE, a large city in the Northeast, the FT group consists of inner city Black and Puerto Rican children. Much of the parent interview data are missing, making it impossible to compare the SES and family structure characteristics of the FT and NFT groups. The FT group has a higher percentage of children with preschool experience (FT = 60%; NFT = 30%) and shows no differences on the kindergarten entry level tests.

At site W, 82% of the FT children have Spanish surnames, compared with 33% of the NFT group. The FT group also had a lower level of family income than the NFT group, but a somewhat higher percentage of families with both a mother and father present in the home (FT = 91%; NFT = 78%). These differences indicate that the FT group consists for the most part of poor but stable Mexican-American families and that the NFT group is a somewhat more affluent majority group.

Site S is different from both other sites in that the non-English speaking group is Cajun French. The community in which the FT schools are located is composed of 82% White Cajun French speakers, and Black English speakers account for most of the rest. The NFT schools have 74% Black English speaking and 26% White Cajun French speaking pupils. The FT children are 52% Black English speaking and 48% White Cajun French speaking. Otherwise, the two groups are closely matched on all indicators, including the percent of children who attended perschool (FT = 75%; NFT = 68%).

TABLE VI-35

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR SEDL Site NE

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	4.64	1.45	14	missing		
Father's Education	missing			missing		
Income	10.07	6.50	14	missing		
Occupation	15.58	1.08	12	missing		
Family Structure						
Mother and Father Present	43%	-	6	missing		
Mother, not Father, Present	57	-	8	missing		
Other	0	-	0	missing		
TOTAL			14	missing		
Entry Level						
WRAT	24.13	8.87	36	21.00	9.48	18
PPVT	31.76	18.18	36	31.88	16.62	18
Preschool Experience						
Attended Headstart or other Preschool	86%	-	55	30%	-	10
TOTAL			64			33
Ethnicity						
Black	50%	-	18	76%	-	14
White	5	-	2	12	-	2
Spanish	45	-	16	12	-	2
Other	0	-	0	0	-	0
TOTAL			36			18

TABLE VI-36

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

SEDL Site W

	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	4.35	1.63	43	4.76	1.52	53
Father's Education	4.33	1.71	39	4.73	1.81	44
Income	8.90	6.15	42	10.30	6.82	53
Occupation	16.40	2.16	43	16.40	2.23	52
Family Structure						
Mother and Father Present	91%	-	39	78%	-	41
Mother, not Father, Present	9	-	4	13	-	7
Other	0	-	0	9	-	5
TOTAL			43			53
Entry Level						
WRAT	21.85	11.91	40	28.00	9.43	40
PPVT	38.58	15.50	40	47.13	11.02	40
Preschool Experience						
Attended Headstart or other Preschool	22%	-	9	12%	-	6
TOTAL			41			49
Ethnicity						
Black	3	-	1	3%	-	1
White	10	-	4	64	-	26
Spanish	82	-	33	33	-	13
Other	5	-	2	0	-	0
TOTAL			40			40

TABLE VI-37

PUPIL AND FAMILY BACKGROUND CHARACTERISTICS FOR

SEDL Site S

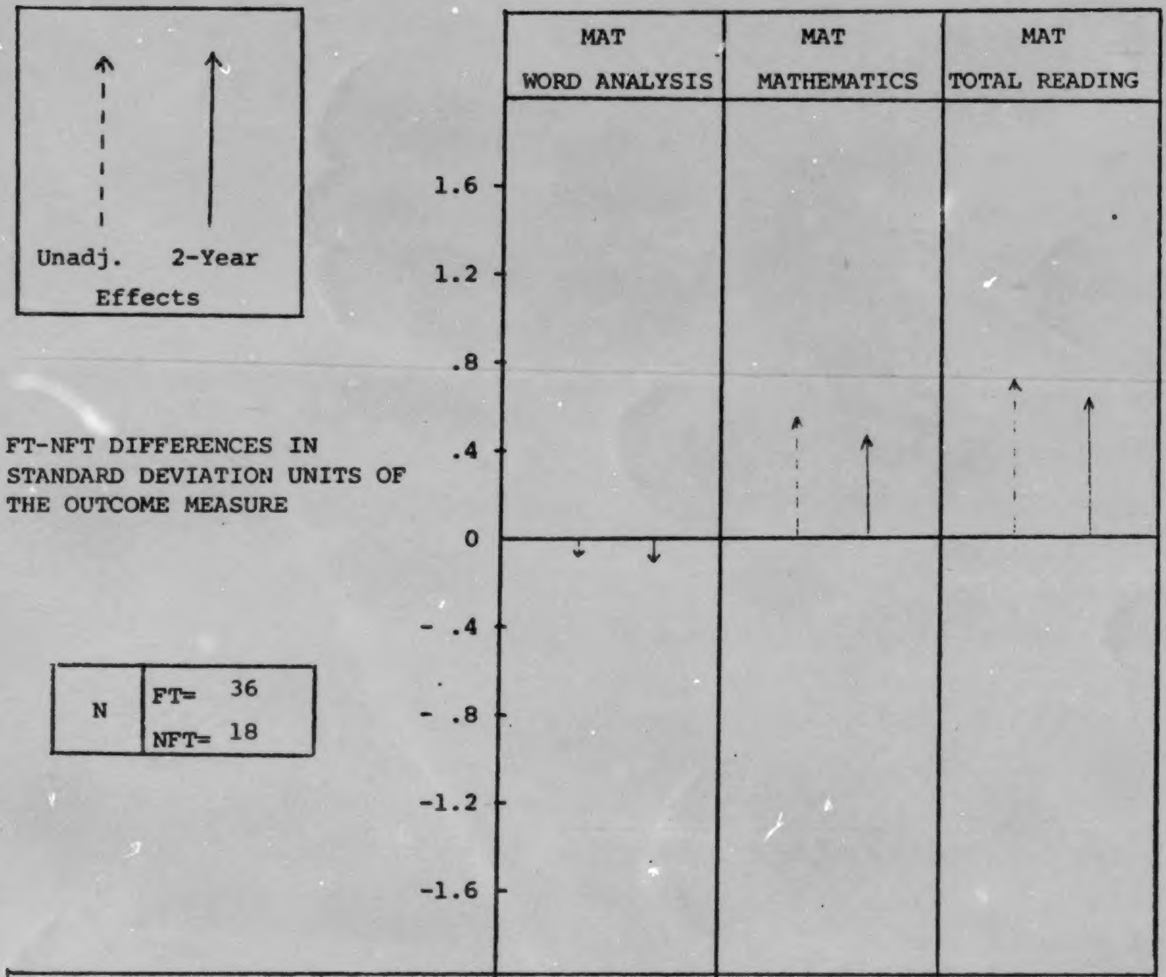
	FT			NFT		
	$\bar{X}$ or %	SD	N	$\bar{X}$ or %	SD	N
Socioeconomic Status						
Mother's Education	4.60	1.41	83	4.68	1.25	34
Father's Education	3.96	1.76	70	4.18	1.81	28
Income	7.08	6.78	84	7.43	6.74	35
Occupation	15.56	2.23	80	16.27	2.18	33
Family Structure						
Mother and Father Present	82%	-	70	80%	-	28
Mother, not Father, Present	15	-	13	17	-	6
Other	3	-	2	3	-	1
TOTAL			85			35
Entry Level						
WRAT	23.40	9.75	64	21.22	7.38	26
PPVT	37.60	14.32	64	35.44	11.02	26
Preschool Experience						
Attended Headstart or other Preschool	75%	-	59	68%	-	17
TOTAL			79			25
Ethnicity						
Black	52%	-	33	74%	-	19
White	48	-	31	26	-	7
Spanish	0	-	0	0	-	0
Other	0	-	0	0	-	0
TOTAL			64			26

6.10.3 By Site 2-Year Effects

The two sites serving Spanish speaking children (NE, Figure VI-52; W, Figure VI-53) both show very strong FT effects on the Mathematics and Total Reading subtests. Site W also shows strong effects in Word Analysis. Although these two sites both serve Spanish speaking children, the sites are located at opposite ends of the nation, and the children come from very different ethnic groups (Puerto Rican and Mexican-American). The SEDL program appears to be working quite effectively with these segments of the bilingual population of the country.

The third site (S, Figure VI-54) shows that the FT group containing both black and Cajun French is not, in the aggregate, doing as well as the other two sites. There is a trend toward a Mathematics effect, but this is mentioned only because it represents a dramatic shift from the performance of this group in mathematics in kindergarten (this longitudinal issue is explored in Monograph 1 of Volume IIB of this report). It is not clear whether there is a differential effect among the non-English speaking and the native English speakers in the FT group which might occur in a bilingual program. In any case, the SEDL program is having a strong impact on the reading development of Spanish speaking children at the end of first grade, but it is associated with a negative effect in reading in a mixed black and Cajun group. Clearly there is much of importance in the SEDL program and a number of problems in bilingual reading instruction still to be solved.

FIGURE VI-52 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
SEDL SITE NE

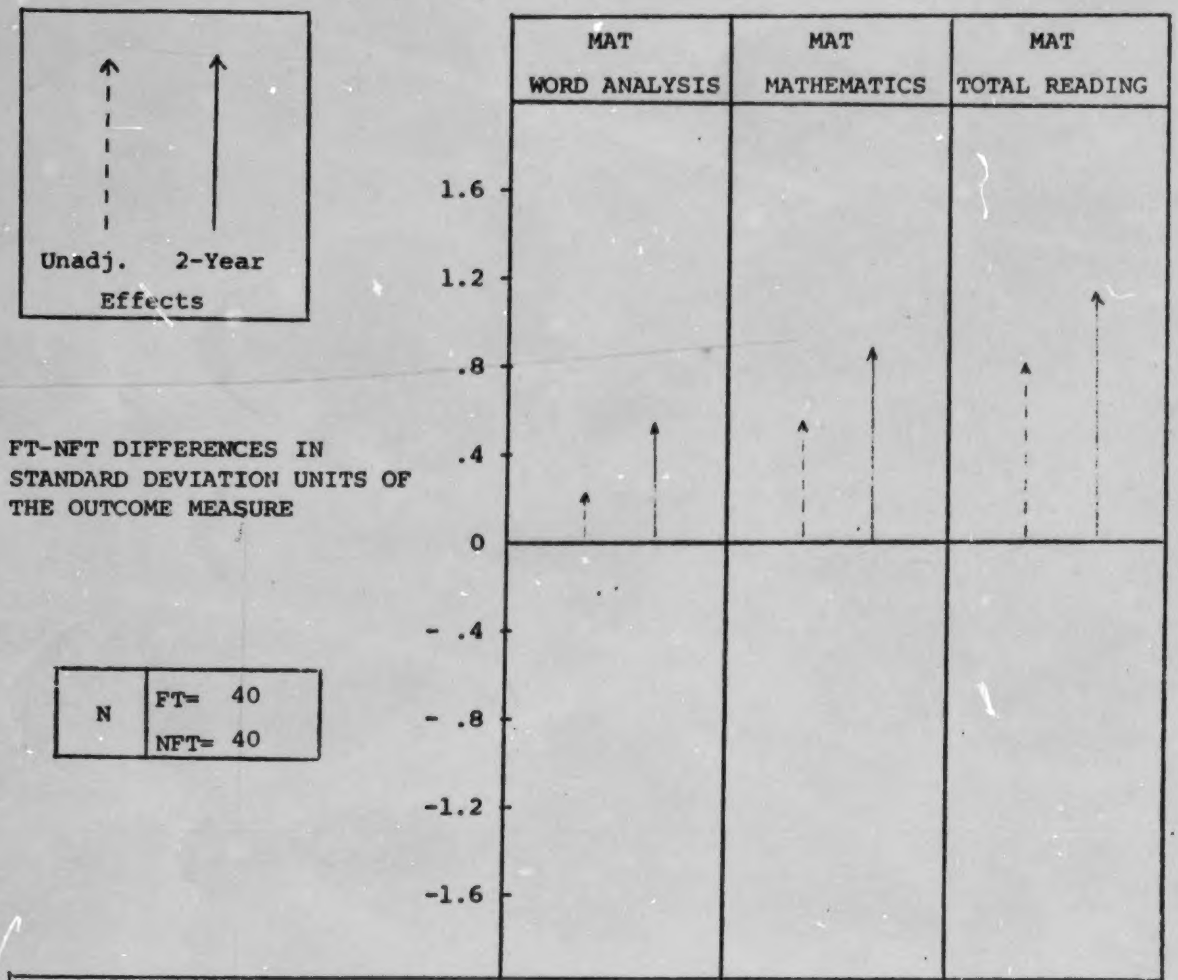


N	
	FT= 36
	NFT= 18

		MAT WORD ANALYSIS	MAT MATHEMATICS	MAT TOTAL READING
B = Sponsor Effect (FT-NFT Differences)	Unadjusted	- .018	8.385	9.812
	2-Year	- .027	6.971	8.948
S.E. = Standard Error of B	Unadjusted	.074	4.455	3.970
	2-Year	.078	4.489	4.117
F = Significance Statistic	Unadjusted	.059	3.542	6.111*
	2-Year	.120	2.412	4.722*
SD=Raw Score Standard Deviation		.247	15.303	13.948
B/SD = Effect in SD Units (See Graph)	Unadjusted	- .073	.548	.703
	2-Year	- .109	.456	.642

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p $\leq$ .05

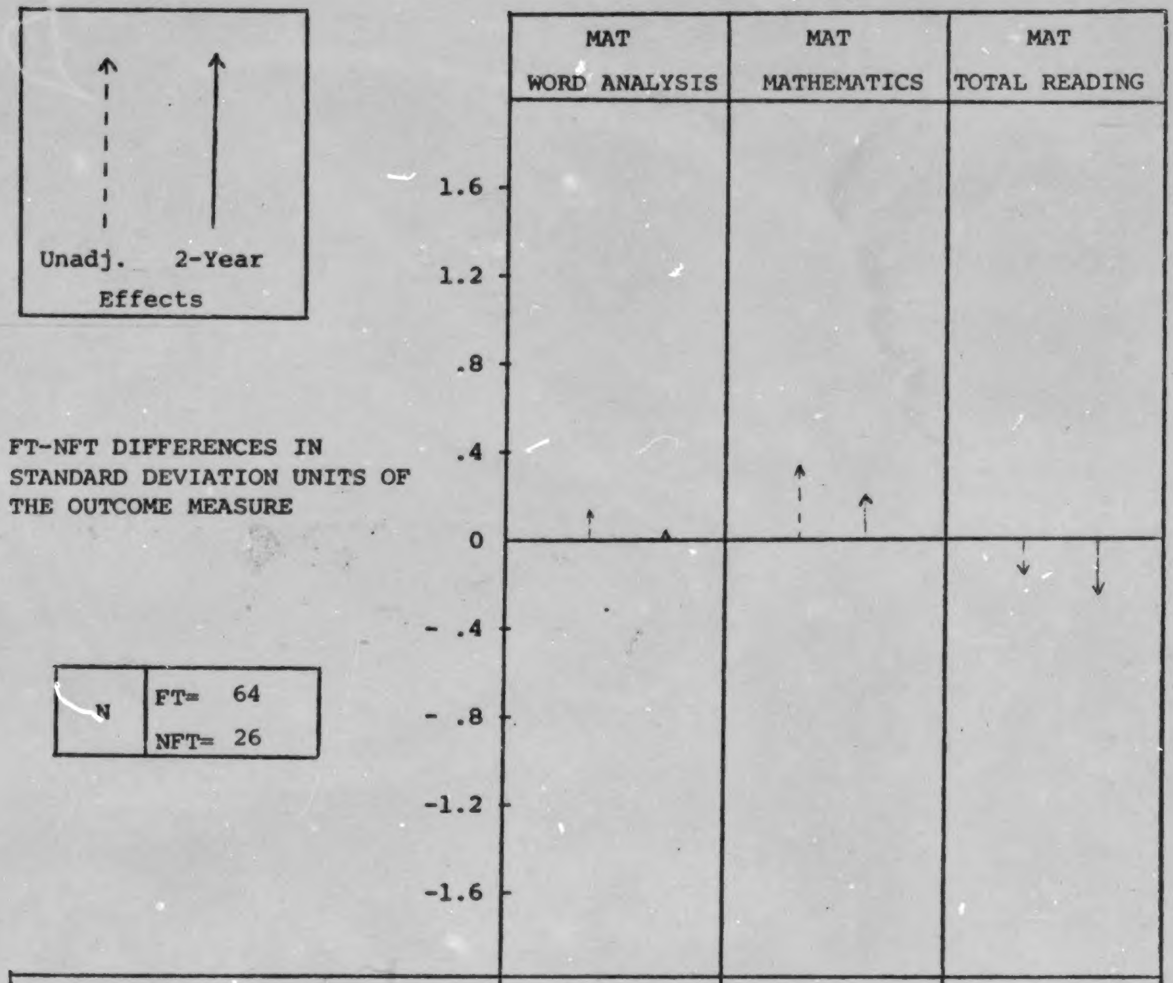
FIGURE VI-53 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
SEDL SITE W



B = Sponsor Effect	Unadjusted	.039	5.900	9.575
(FT-NFT Differences)	2-Year	.100	9.096	13.326
S.E. = Standard	Unadjusted	.041	2.268	2.402
Error of B	2-Year	.036	1.820	2.062
F = Significance	Unadjusted	.904	6.765*	15.888*
Statistic	2-Year	7.717*	24.980*	41.770*
SD=Raw Score Standard Deviation		.183	10.505	11.710
B/SD = Effect in SD	Unadjusted	.213	.562	.818
Units (See Graph)	2-Year	.546	.866	1.138

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates p ≤ .05

FIGURE VI-54 END OF FIRST GRADE FT/NFT CONTRASTS
TRACED 2-YEAR SAMPLE
SEDL SITE S



B = Sponsor Effect (FT-NFT Differences)	Unadjusted	.022	4.030	- 1.535
	2-Year	.003	2.416	- 2.565
S.E. = Standard Error of B	Unadjusted	.045	2.804	2.257
	2-Year	.039	2.268	1.990
F = Significance Statistic	Unadjusted	.239	2.065	.462
	2-Year	.006	1.134	1.662
SD=Raw Score Standard Deviation		.194	12.197	9.736
B/SD = Effect in SD Units (See Graph)	Unadjusted	.113	.330	- .158
	2-Year	.016	.198	- .264

NOTE: 2-Year values are adjusted for kindergarten entry score and mother's education
Blank columns indicate heterogeneity of regression and are not reported here
An asterisk indicates $p \leq .05$

CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

The major focus of the analyses presented in this report is on the effects of two years of Follow Through on children who remained with the same Sponsors over that time period. It is important to note that this two year period includes two distinct educational treatments: a kindergarten year and a first grade year. Although these different years may have produced very different contributions to the full two year effect (indeed, as is reported in Monograph I, Volume II-B of this report, for some Sponsors the contributions were in opposite directions), the purpose of the present report is to describe the status of FT children at the end of first grade as a function of two years of participation in the program. The fact that the end-of-first grade status could have been preceded by a much higher status, a much lower status, or an equal status at the end of kindergarten may be an educationally interesting finding, but it does not focus on the question of major interest of this section of the report. Here we wish to know the total accumulated effect of two years of consistent experience in a Follow Through program. For this reason, the only adjustments for initial differences between FT and NFT were made on differences observed upon entry into kindergarten.

Two major points need to be summarized:

- The children involved with three very different Sponsors show important contrasts with their NFT comparison groups;
- For most of the Sponsors, site dynamics appear to mediate or even to attenuate their effects.

Each of these points is discussed in turn below.

First, it is clear that three very different programs are contributing importantly to their FT children's academic growth. The University of Kansas, the Educational Development Center, and Southwest Educational Development

Laboratory are each showing important mathematics effects on the average over all sites with which they are working and in some domains in some sites. Each Sponsor's effects vary to some extent from site to site, but the aggregated effects over all sites with which they are working are large enough in some domains to report that at the end of first grade they have produced consistent and robust effects in these domains.

Of the remaining Sponsors, three other very diverse Sponsors are apparently producing some important contrasts, but these are so variable over sites that when all children from all sites for any one of these three are analyzed as a single FT/NFT comparison, there are no apparent effects of consequence. These three are the Far West Educational Laboratory, the University of Oregon, and the University of Florida.

The remaining four Sponsors typically show either no important contribution over the NFT experience or significantly smaller effects than the NFT. These Sponsors are the University of Arizona, Bank Street College of Education, High Scope Foundation, and the University of Pittsburgh. It is also true that the findings for these Sponsors are those in which we have the least confidence. That is, in each case, there is good reason to suspect the meaningfulness of the FT/NFT comparisons and to argue that the findings are not useful estimates of the power of the educational programs involved.

First consider the Sponsors for whom strong effects are apparent.

1) University of Kansas. Overall, this Sponsor shows strong effects on Word Analysis and Mathematics, and a trend toward an FT effect on Total Reading. This pattern is obvious and magnified in three of the sites; it is almost exactly the opposite in the fourth site. The differences between these two groups of sites are very important leads to future work in understanding the basis of the impact of the Kansas program. The first three sites all involve very low income Black families whose Follow Through children entered kindergarten with generally low academic scores, but not any lower than those produced by their NFT counterparts. The NFT children come from essentially the same communities although their families tend to have somewhat higher incomes than the FT families. The communities served are all urban, and the Kansas program appears to have made a major academic impact where so many other anti-poverty, inner city oriented programs have failed.

The situation is quite different in the fourth site where the community is rural, White, and not particularly poor. The FT children reflect their community in these respects, and, in fact, far exceed their NFT comparisons on the entry level test scores. Despite these advantages, the NFT children are significantly ahead of the FT children on Total Reading and Word Analysis and slightly ahead of them on Mathematics at the end of first grade. The Kansas program appears to have differential effects depending upon the ethnic group and/or the urban rural character of the site. It is certainly premature to attempt to account for ethnic differences in response to a token reinforcement program or rural/urban differences in response to the program. It is not premature, on the other hand, to decide that this is a valuable topic to explore and that a strong focus in the future on the nature of the sites and the characteristics of the children involved with this Sponsor would be of major benefit to both the FT national evaluation and the Sponsor.

2) Educational Development Center. Overall, this version of the British Infant School model shows strong effects on Mathematics, no effects on Total Reading (i.e., the FT and NFT children are performing at about the same adjusted level), and negative effects on Word Analysis. It is worth noting that this Sponsor has one of those innovative educational approaches which has not only redesigned the mode of instruction, but has also restructured the curriculum to a very large extent. Unlike the token reinforcement approach of Kansas in which the traditional curriculum is administered in a manner that allows for systematic reinforcement of appropriate behaviors, the EDC model has excluded many traditional substantive materials from the classroom, in accordance with an alternative notion of the kinds of materials children need to facilitate intellectual growth. EDC's curriculum aims at some objectives which are not directly measured by the Follow Through test battery.

On the other hand, number manipulation and mathematics concepts are definitely topics of major concern, although not in the traditional mode of presentation. It is reasonable to expect that a traditional mathematics test such as the MAT will provide tasks which are at least within the realm of skills acquired through the EDC mode. Word Analysis probably does not tap any skills which are closely related to those being nurtured in the EDC classrooms.

The site variability for this Sponsor is somewhat more revealing than the overall results. The only effects found consistently across sites are those for mathematics. The remaining outcomes vary across sites. One large city site serving a predominantly Black population whose children score low at entry into kindergarten, has exceptionally strong first grade effects on all outcome measures. A second site serves a small town, but with a large enough Black population to fill up the FT classes. Here, along with strong positive Mathematics effects, there are strong negative effects on Total Reading and Word Analysis. In the third site, which is almost completely White, a Mathematics effect is apparent, along with very small, non-significant trends in favor of FT in the remaining outcome measures.

There is variability here in both the characteristics of the sites and in the nature of the findings, but it is, once again, premature to attempt to account for these findings in terms of the nature of the educational model being implemented. It is clear, however, that at least in Mathematics, the EDC approach has a consistent impact upon a variety of kinds of children in a variety of settings. This is compelling evidence that the instructional mode can account for a major share of the effects. On the other hand, each of these sites provided the great majority of children with preschool experiences; in at least one of these sites, moreover, the preschool experience was organized and administered by EDC as a downward extension of their model. It is in this site that the largest effects are found.

It seems at one level, therefore, that a necessary condition for the emergence of effects in EDC children may be a year of preschool experience. This seems reasonable if it can be assumed that preparation for EDC's very non-traditional school activities is a necessary prerequisite to successful interaction with the EDC model. The assumption is based on the notion that both children and teachers need a good deal of time getting used to exploring the world without prejudice, fear of failure, and with a sense that the skills acquired thereby can make a difference in meeting the challenges of school work. If this intellectual socialization process is in fact required for the observed effects at the end of first grade, it would be reasonable to expect that the kindergarten results would be less strong than the first grade effects because kindergarten is a time of continuation of the socialization rather than a time of utilization of skills. The comparison of the kindergarten and first grade effects in a longitudinal mode is the focus of Monograph I, Volume II-B of this report, and will not be discussed here. It is worth noting, however, that just such diminished kindergarten effects have been observed, and the first grade effects discussed in this chapter follow directly from those early experiences. The EDC model is complex and changing and clearly needs to be examined in detail as the data for future grades become available.

3) Southwest Educational Development Laboratory. This Sponsor was the developer and administrator of bilingual educational programs. As such, the sites with which the SEDL FT programs are involved all include large numbers of non-English speaking children. In this case, the three sites include a Mexican-American site on the West Coast, a Puerto Rican site in a large eastern city, and a Cajun French site in the south.

The overall analysis of the data collected from all three sites is not very revealing because two of three outcomes yielded heterogeneous regressions with the WRAT kindergarten pretest. In both cases it is apparent that FT children from the upper segment of the WRAT distribution had much higher scores on the outcome measures than the NFT children from the upper end of the WRAT distribution. The SEDL program appears to have a differential impact in children depending upon their pretest scores, and since these heterogeneities disappeared when the data were analyzed at site level, it is clear that sites differ in their effects with respect to the kinds of children who live there. This point will be explored in the next section, but it is important to note that the SEDL program produced very strong effects on the two heterogeneous outcomes (Mathematics and Total Reading) in the two Spanish speaking sites and weak or negative effects in the Cajun site. Since all three FT groups had essentially the same WRAT pretest scores, the heterogeneity findings here are a clear example of site differences defined in terms of native language differences.

The site level effects reveal one other point about the SEDL program. The Puerto Rican and Mexican-American sites both showed very strong effects on Mathematics and Total Reading, whereas the Cajun site showed its only effects (albeit small and non-significant) on Mathematics. Total Reading showed negative effects in the Cajun site, a fact which may reflect the totally non-English context within which these children live as much as the problems in the SEDL French/English curriculum. It is not possible to conclude that SEDL produces effects in the relatively non-verbal domain of Mathematics in the full range of cultural groups, but is thus far effective in reading instruction only when dealing with Spanish speaking groups. Apparently, teachers skilled in the bilingual mode of instruction have developed a sensitivity to the needs of children who find themselves in a school system dominated by a culture different from their own. The need to have the children's culture respected, to have the children's difficulty with English treated sensitively by those who can speak the children's native language might be reasonably met by those trained in the bilingual mode. Such teachers may be of great support to any group of non-English speaking children learning mathematics regardless of what the children's native language is. This could be true even in those situations (e.g., Cajun) in which the language instruction is not fully worked out.

In any case, this Sponsor, EDC, and Kansas are quite effective at the end of first grade. These are three very different approaches to compensatory education, and in general these three are dealing with quite different samples of children. Comparison among these three is not, under the conditions of this study, possible. The best that can be stated as summary is that there are several routes to effective education for low income children, and these routes may be specific to places and types of children. In order to see this point in better perspective, we should now consider the remaining Sponsor effects.

The second group of Sponsors includes those who are producing some important effects which are so variable across sites that the findings are obscured when the data are aggregated to overall Sponsor level. These Sponsors are Far West, Oregon and Florida.

1) Far West Laboratory. This Sponsor, offering a responsive environment model, is working with sites that are either predominantly Black, predominantly White, or mixed Black and White, Anglo and Spanish speaking, and are all located in different parts of the country. Such variability of places and people would ordinarily provide the beginning of an opportunity to study the robustness of the model under different conditions. However, these differences are also associated with a great deal of variability in the extent to which the FT and NFT groups are mismatched on entry level measures. There is no discernible pattern to the effects except that where the mismatch is smaller the effects are larger. However, the mismatches are themselves so highly confounded with place and ethnic variability, that it is almost impossible to account for the variability in effects over sites. Far West has the full range of effects patterns from strongly positive effects on all three outcomes to strongly negative effects on all three outcomes. These are important findings in one sense since they indicate that there are site conditions which are associated with the realization and the restriction of the potentials of this model. There are conditions, that is, which allow the model to work and conditions which keep it from working, and until a detailed study of the operation of the Far West programs at selected sites is carried out, much of importance to the compensatory education literature is kept buried.

2) University of Oregon. This reinforcement-oriented program showed very impressive effects in kindergarten and strikingly diminished results after two years. Site level effects are variable and in some cases strongly positive: All three sites with which this Sponsor is working are predominantly Black, although they are located in different parts of the country. One very important difference between them is that in two sites the FT groups have lower SES than the NFT groups, but higher entry level scores, and the third site has children from lower SES families

but who are low achievers (relative to NFT) on the entry tests. This last site yields low positive effects for FT; the other sites tend to show strong effects in reading (but not in Mathematics). Apparently Oregon is effective in those sites in which the FT children have an advantage when they enter the program. It is not clear whether the stronger effects are the result of the inability of the covariance technique to adjust for these kinds of initial differences, or whether, on the other hand, some other site factor is associated with the selection of higher achieving children from among the lower income groups for participation in FT. If there is a selection factor which reflects some site-specific factor, then this factor may contribute other child properties which account for the strong effects. These are issues which must be considered before we can conclude that Oregon is effective with higher achieving low income children, although the data do seem to point toward such a conclusion at this point in the study.

3) University of Florida. This parent education program, which also falls in the cognitive stimulation category and is educationally very different from the Oregon program, shows a pattern of effects which is rather similar to the Oregon pattern. The Florida program is working with very economically disadvantaged children, although some are White and some are Black. At three of the four sites, the FT children, while from very low income families, score higher on the entry tests than the NFT group. These are the sites which show the stronger effects, particularly in reading. The fourth site also involves low income White children, but these children score well below their NFT comparisons on entry tests. It is in this fourth site that all of the effects are negative. Much the same can be said about the confounding of site and mismatch in the Florida program as was said about the Oregon program. Certain conditions seem to facilitate very impressive effects for this Sponsor, but these conditions seem to be site-specific and must, therefore, be sought at site level. Until this is accomplished, important insights about the working of the Florida parent education program will remain uncovered.

The remaining Sponsors, University of Arizona, Bank Street, High Scope, and University of Pittsburgh, all show very few positive effects and a relatively large number of negative effects. In each case, however, serious problems of interpretation result from severe mismatch or major sampling problems. It would be difficult to draw any conclusions about these Sponsors except to say that these problems are obviously associated with a variety of educational models. We do not seem to have here so much an opportunity to examine issues related to educational theory as an opportunity to examine issues of implementation. Once again, the focus should be on the individual site, since very different events seem to be happening at the various sites with which these Sponsors are involved and even though the effects patterns appear to be uniformly poor. Only at the site level can we determine the full reasons for this pattern. This should be the focus for future work.

The issue of site variability has been documented in the discussion above and requires little comment at this point. What is important is that the site variations not only make each Sponsor's effects difficult to interpret, but in some cases they also impose statistical restrictions on the interpretations, especially where heterogeneity of regressions has been reported. Sites appear to differ in the nature of the FT/NFT mismatches, so that in one site the FT children are ahead of the NFT children on entry tests and in another site the NFT children are ahead. In other sites the mismatch is in the ethnic composition of the FT and NFT groups. At the same time, Sponsors seem to be uniquely effective or ineffective with one or the other of the mismatched groups, so that one kind of relationship between a covariable (the mismatch variable) and an outcome is produced in the FT group and another kind of relationship between these variables is established in the NFT group. This is the heterogeneous relationship which precludes covariance analyses.

It seems clear that mismatches do not occur by chance, but rather that they reflect unique site properties. There are reasons why disproportionate numbers of high or low achieving children are assigned to the FT or the NFT groups. There are reasons why the FT group is made up of one ethnic group and the NFT of another. These reasons reflect a great range of community dynamics and undoubtedly influence the kind of operational relations each Sponsor establishes with the local school system. The mismatch thus represents a confounding of child properties and community properties. The problem can only be unraveled by detailed studies at the site level. Each Sponsor must approach its own sites in a unique way, and this can only be understood as the socioeducational dynamics of the mismatch are revealed.

Focussing on site studies should not pose a problem for the federal policy maker who is interested in establishing a federal approach to compensatory education. A major result of site studies can be a directory of effects associated with types of Sponsors under types of site conditions. This could be a very useful document for the local decision maker who needs to estimate the consequences of educational choices using the maximum amount of information. A directory of the sort described here would speak directly to those needs and would put the federal government directly in the position of supplying locally useful information based on national studies. At this time in the history of the national Follow Through evaluation, still two years away from the full four year longitudinal analysis, the only step in decision making which the federal government can take which would not be premature is to start assembling just this kind of directory. Further analyses would serve to update, clarify, and make more precise the entries in the directory which would in the meantime have increasing value to local educators. Along with a continuation of the Follow Through evaluation to its logical and planned conclusion, the starting of this directory is the major recommendation of this second annual evaluation of Follow Through.

APPENDICES

APPENDIX A

THE ANALYTIC SAMPLE

APPENDIX A

THE ANALYTIC SAMPLE

1.0 INTRODUCTION

This Appendix presents a description of the sample upon which our analyses of covariance is based. This sample is a unique selection of children and is not representative of either the Follow Through population in particular or disadvantaged children in general, although it has many characteristics in common with both of these populations. This is, however, the sample which gives rise to our findings and a detailed description is essential if we are to interpret and understand the results. Such a description will also allow the reader to draw generalizations on an ad hoc basis.

In order to fully describe the sample we will focus on several facets of the pupils and their home/school/community environment. The pupil characteristics chosen for descriptive purposes are age, sex, ethnicity, preschool experience, and entering achievement. This last characteristic is measured in the Fall of the kindergarten year by the Wide Range Achievement Test (WRAT) and the Peabody Picture Vocabulary Test (PPVT). In addition, a set of family background variables has been included to reflect the socio-cultural environment of the pupil's home. Developed from responses to a Parent Interview administered in the Fall of 1971, these variables include family income, educational level of the child's mother, number of people in household, and structure of the family (mother and father both present, mother but not father present, or other). Two parent attitudinal variables are also presented: parental Locus of Control or feelings of efficacy over life in general and parental Perception of the School's Receptivity to their feelings and ideas.

The context in which FT programs operate involves ever widening spheres. The most immediate context is the classroom, where the pupil is exposed to his teacher and his peers. To reflect that environment, we have included teacher education, teacher experience, and teacher ethnicity as descriptors of the primary adult in the classroom.[1] We have also included the degree of ethnic integration of the class and the entry level of the class to describe the characteristics of the pupils' peers.

The next sphere of influence is the school. Here we follow Coleman's (1968) distinctions between two kinds of input resources. First are the characteristics of the pupils in the school--or what Herriott (1973) describes as throughputs. These are elements over which school administrators can exercise relatively little control (except in such cases where district or school boundaries, building construction, and feeder school patterns are deliberately modified to influence school composition). To describe the throughputs of the schools in the analytic subset, we have included the ethnic composition of the pupils, as reported in the U.S. Office of Civil Rights Directory, 1970.[2]

We would have liked to include in this section school mean entry level, income and mother's education as other measures of school throughputs. However, the only data available for these indices were those based on an aggregation of the children in our sample and not representative of the total school throughput. In this descriptive discussion of the sample these means did not add any information beyond that presented by the non-aggregated child measures of entry level, income and mother's education. Hence they will not be discussed.

The other school variables we have included are those which Herriott has labelled inputs--things over which school administrators have some control.[3] These include the amount of auxiliary services being delivered to FT/NFT children in the school, the amount of Title I dollars per pupil in the school, the student-teacher ratio of the school, and school enrollment.

---

[1] For a more complete description of the teachers, the reader is referred to Monograph II, "Teachers".

[2] Office of Civil Rights Director of Public Elementary and Secondary Schools in Selected Districts: Enrollment and Staff by Racial Ethnic Group, U.S. Government Printing Office, 1970. It should be noted that this variable is independent of the FT sample and represents a context factor.

[3] We shall see in the Monograph III, "School and Community", that the relationship of inputs in throughputs is intertwined. With the advent of large scale social welfare and education programs for the disadvantaged, many additional inputs, purchases by Title I and other federal funds, are found in the lower SES schools.

Finally, the context in which education takes place includes the greater community. To represent community, we have selected two variables which in the past have been found related to both educational inputs and outputs: region and city size. The reader is referred to Monograph III for a more detailed description of the demographic characteristics of the communities operating FT programs.

2.0 PUPIL CHARACTERISTICS

Turning to pupil characteristics, we find that, overall, FT pupils in the analytic subset are younger than NFT pupils upon entry to kindergarten. As seen in Table A-1 (age in months as of December 31, 1972), FT children are on the average about one-half month younger than their counterparts--a difference significant at the .01 probability level. This age differential is consistent across all Sponsors, with the exception of High Scope. For the University of Kansas, EDC and SEDL, the difference is both significant, and sizeable, exceeding one month in each case. Whether or not the age disparity reflects a conscious decision on the part of school administrators to provide early education to the most needy children (in the form of age waivers) or the withholding of children from school by NFT parents is difficult to say.

As can be seen in Table A-2, the FT group has slightly more girls in it (53.2%) than the NFT group (50.1%). This small differential is consistent for seven out of the ten Sponsors, with only High Scope, Pittsburgh and SEDL having slightly more boys than girls. It is possible that the finding that there are more girls in the FT group than in the NFT group is closely akin to the FT/NFT age differential. Parents may tend to keep boys at home somewhat longer than girls at this age when girls are apt to mature at a faster rate.

Turning to ethnicity, Table A-3 presents the percentage of White children in the analytic sample by Sponsor FT/NFT group. Approximately one-half of the NFT group is White, whereas approximately one-third of the FT group is White. The large number of minority children found in both groups reflects the populations to be found in largely urban, Northern communities from which many of the children in this analytical sample are drawn. Moreover, the fact that there are substantially more minority children in the FT program, which is targeted for children from low income families, is consistent with the depressed socioeconomic status of such groups in these communities.

TABLE A-1

MEANS AND STANDARD DEVIATIONS OF CHILD'S AGE IN MONTHS BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.
FAR WEST LABS	178	69.719	4.17	219	68.968	4.00
UNIVERSITY OF ARIZONA	166	69.361	4.12	260	69.096	4.15
BANK STREET	67	67.463	5.09	119	67.084	4.56
UNIVERSITY OF OREGON	82	67.890	4.19	95	67.347	3.68
UNIVERSITY OF KANSAS	113	68.973	6.31	320	67.366	3.72
HIGH SCOPE FOUNDATION	145	67.931	4.11	168	68.786	3.84
UNIVERSITY OF FLORIDA	74	69.797	6.04	173	69.405	4.20
EDUCATIONAL DEVELOPMENT	96	67.646	3.99	150	66.333	4.81
UNIVERSITY OF PITTSBURGH	60	69.900	4.23	153	69.719	4.81
SEDL	50	68.400	3.87	67	66.925	3.51
ALL SPONSORS	1031	68.795	4.69	1724	68.255	4.27

Data source: Follow Through Class Roster

TABLE A-2

SEX OF CHILD BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	MALE	FEMALE	N	MALE	FEMALE
FAR WEST LABS	178	50.6%	49.4%	219	44.7%	55.3%
UNIVERSITY OF ARIZONA	166	56.6	43.4	260	48.8	51.2
BANK STREET	67	56.7	43.3	119	50.4	49.6
UNIVERSITY OF OREGON	82	47.6	52.4	95	42.1	57.9
UNIVERSITY OF KANSAS	113	46.0	54.0	320	43.4	56.6
HIGH SCOPE FOUNDATION	145	50.3	49.7	168	53.6	46.4
UNIVERSITY OF FLORIDA	71	50.7	49.3	173	48.6	51.4
EDUCATIONAL DEVELOPMENT	96	45.8	54.2	150	42.0	58.0
UNIVERSITY OF PITTSBURGH	60	41.7	58.3	153	49.0	51.0
SEDL	50	44.0	56.0	67	46.3	53.7
ALL SPONSORS	1028	49.9	50.1	1724	46.8	53.2

Data source: Follow Through Class Roster

TABLE A-3

STUDENT ETHNICITY BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	Minority	White	N	Minority	White
FAR WEST LABS	178	43.8%	56.2%	219	55.7%	44.3%
UNIVERSITY OF ARIZONA	166	31.3	68.7	260	62.7	37.3
BANK STREET	67	9.0	91.0	119	33.6	66.4
UNIVERSITY OF OREGON	85	80.5	19.5	95	94.7	5.3
UNIVERSITY OF KANSAS	113	79.6	20.4	319	88.1	11.9
HIGH SCOPE FOUNDATION	145	63.4	36.6	167	84.4	15.6
UNIVERSITY OF FLORIDA	74	29.7	70.3	173	67.1	32.9
EDUCATIONAL DEVELOPMENT	96	37.5	62.5	150	78.0	22.0
UNIVERSITY OF PITTSBURGH	60	1.7	98.3	153	1.3	98.7
SEDL	50	52.0	48.0	67	50.7	49.3
ALL SPONSORS	1031	45.5	54.5	1722	64.2	35.8

Data source: Follow Through Class Roster

These ethnic differences are not, however, consistent within Sponsor groups. While it is true that for eight out of ten Sponsors, the FT group has more minority children than the NFT group, these comparisons belie very great differences among the eight Sponsors. University of Oregon, University of Kansas, and High Schope are represented primarily by minority children as are their NFT counterparts. Both FT and NFT groups for Far West Labs and SEDL are approximately 50 percent minority. University of Arizona, University of Florida and EDC have substantially more minority children in their FT groups than in their NFT groups. For Bank Street, the NFT group is over 90 percent White; the FT group is two-thirds White. And finally, the University of Pittsburgh FT and NFT groups are both more than 98 percent White. It is clear that, for a variety of reasons, systematic differences appear in the ethnic composition of the analytic sample which make Sponsor to Sponsor comparisons difficult.

Table A-4 presents the percent of children with preschool experience by Sponsor FT/NFT group. The sample size (N) in this table reflects the missing data on this variable. Apparently many parents and teachers are unable to accurately report a child's preschool experience at the time the child enters kindergarten. This is a very important variable for which an appropriate information system has not yet been developed. Seventy-one percent of the children in the FT group for whom data are available have attended Head Start and/or an equivalent preschool program whereas only 34 percent of the NFT group have done so. It appears that overall FT is achieving its stated objective of reaching children with preschool experience. While FT children in all Sponsor groups have had more preschool experience than their NFT counterparts, differences among Sponsors are great. In the FT group, for example, the range of percent of children with preschool experience is from 100 percent for the University of Oregon to 58 percent for the University of Florida. In the NFT group, the range is from 71 percent for the University of Kansas group to 11 percent for the University of Florida. If, as the Follow Through plan implies, preschool is a source of academic growth upon which Follow Through must build, Sponsors vary greatly as to the availability of this base for their building program.

TABLE A-4

PREVIOUS PRESCHOOL EXPERIENCE BY SPONSOR BY FT/NFT

	NFT			FT		
	N	No	Yes	N	No	Yes
FAR WEST LABS	150	72.7%	27.3%	194	22.2%	77.8%
UNIVERSITY OF ARIZONA	89	79.8	20.2	237	38.8	61.2
BANK STREET	30	63.3	36.7	92	39.1	60.9
UNIVERSITY OF OREGON	54	77.8	22.2	44	0	100.0
UNIVERSITY OF KANSAS	93	29.0	71.0	305	31.5	68.5
HIGH SCOPE FOUNDATION	62	43.5	56.5	148	16.2	83.8
UNIVERSITY OF FLORIDA	37	89.2	10.8	157	42.0	58.0
EDUCATIONAL DEVELOPMENT	55	72.7	27.3	148	28.4	71.6
UNIVERSITY OF PITTSBURGH	39	69.2	30.8	72	13.9	86.1
SEDL	45	75.6	24.4	62	24.2	75.8
ALL SPONSORS	654	65.6	34.4	1459	29.1	70.9

Data source: Follow Through Class Roster

Finally, we have two measures of pupil entry level: the WRAT and the PPVT, both administered in the Fall of the kindergarten year. Table A-5 displays the mean and standard deviation for the Fall kindergarten WRAT by Sponsor FT/NFT group. The overall FT/NFT groups appear to be fairly well matched. While overall the NFT group is slightly ahead of the FT group on this pretest measure, the difference between the groups is extremely small and not statistically significant.

When we examine the differences among Sponsor FT groups, however, we once again see wide variability. While most Sponsor FT groups cluster around the FT mean, two Sponsors deviate greatly from the others. The University of Pittsburgh appears to be working with a relatively high FT group in terms of entry level; SEDL has a relatively low group.

In respect to FT/NFT differences within each Sponsor, there are three instances of significant mismatch. The Far West Laboratory and the University of Arizona FT groups are well behind their respective NFT groups in the Fall of kindergarten. On the other hand, the University of Kansas FT group is well ahead of its NFT group on this initial measure. It is clear that simple covariance adjustments cannot fully remove these initial differences, or deal effectively with the implications of such differences for growth in learning.

We have found in our analytic sample that the PPVT, more than any other test, is sensitive to family background characteristics. Before turning directly to those characteristics, therefore, we will describe differences in the FT/NFT groups on the Fall PPVT. As can be seen in Table A-6, overall the FT group scores significantly lower than the NFT group on this test. When we examine variability across Sponsors, the range of entry level scores on this measure is not very much different from that on the WRAT. Most Sponsors appear to be represented by fairly similar groups of children, with the exception of University of Pittsburgh and SEDL. These two Sponsors are once again on the high and low ends of the continuum, respectively. FT/NFT contrasts within Sponsor, on the other hand, differ somewhat on the two tests. The University of Kansas and Far West Labs FT/NFT groups, which were both severely mismatched on the Fall WRAT, are not mismatched on the Fall PPVT. EDC and SEDL, which were fairly well matched on the Fall WRAT, have their FT groups significantly below the respective NFT group on the PPVT. Only the University of Arizona appears to be poorly matched on both pretest measures.

TABLE A-5

MEANS AND STANDARD DEVIATIONS OF THE FALL KINDERGARTEN WRAT BY SPONSOR BY FT/NFT

	NFT			FT		
	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.
FAR WEST LABS	178	31.927	11.77	219	28.699	10.87
UNIVERSITY OF ARIZONA	166	35.018	13.86	260	30.058	13.16
BANK STREET	67	26.522	10.69	119	28.319	12.11
UNIVERSITY OF OREGON	82	29.049	12.13	95	30.600	11.17
UNIVERSITY OF KANSAS	113	25.168	11.80	320	28.737	11.51
HIGH SCOPE FOUNDATION	145	28.890	12.24	168	27.250	11.30
UNIVERSITY OF FLORIDA	74	28.351	13.17	173	27.931	12.66
EDUCATIONAL DEVELOPMENT	96	29.771	12.99	150	29.347	12.48
UNIVERSITY OF PITTSBURGH	60	35.417	12.26	153	36.255	12.49
SEDL	50	25.800	9.25	67	23.090	9.51
ALL SPONSORS	1031	30.125	12.69	1724	29.280	12.17

To summarize, important differences exist among FT/NFT groups on a number of key pupil background variables. Overall, the FT children in the analytic sample appear to be younger than their NFT counterparts. There are also more girls and more minority children in the FT than in the NFT group. More children in the FT group have attended preschool. While the two groups are fairly similar in entering WRAT scores, they are dissimilar on entry level as measured by the PPVT. The FT group starts out significantly lower than the NFT group on this pretest measure which appears to be highly related to other home background variables discussed below. Furthermore, there are important differences among Sponsor FT groups and within Sponsor FT/NFT groups on each of these variables. It is apparent that insofar as this analytic sample is concerned Sponsors are working with different types of children and have been more or less well matched with their respective NFT groups.

3.0 FAMILY BACKGROUND CHARACTERISTICS

Tables A-7 through A-10 present the means and standard deviations for four variables: income, mother's education, family structure, and number of people in the household. The first two (together with occupation and housing) have traditionally been used as indicators of SES. The latter two are concommitant variables which may influence the quality of the educational environment in the home.

As can be seen, the FT group is substantially below the NFT group on family income. The FT group has a mean income of roughly $5,600; the NFT group roughly $6,400. While there is considerable variability in income level among Sponsor FT groups, every FT group is below its respective NFT group. For six of the ten Sponsors (Far West, University of Arizona, Bank Street, University of Oregon, High Scope, and University of Florida), moreover, this difference is significant.

In contrast to income, there is no overall difference between the FT/NFT groups on mother's education. Further, there is little variability among Sponsor FT groups on this measure. Only the University of Pittsburgh and SEDL FT groups are deviant from the other Sponsors, being on the high and low ends of the continuum, respectively. There are two significant FT/NFT differences within Sponsor on mother's education. The University of Arizona FT group is significantly below its NFT group.

TABLE A-6

MEANS AND STANDARD DEVIATIONS OF FALL PPVT BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LAB	178	48.421	9.84	219	46.525	11.10
UNIVERSITY OF ARIZONA	166	50.030	12.58	260	43.427	12.45
BANK STREET	67	41.343	15.51	119	43.336	11.42
UNIVERSITY OF OREGON	82	41.610	12.43	95	42.495	10.48
UNIVERSITY OF KANSAS	113	40.319	12.38	320	39.528	11.53
HIGH SCOPE FOUNDATION	145	43.917	13.22	168	42.179	11.81
UNIVERSITY OF FLORIDA	74	46.838	13.04	173	44.861	11.87
EDUCATIONAL DEVELOPMENT	96	46.354	12.36	150	41.300	12.28
UNIVERSITY OF PITTSBURGH	60	54.483	6.24	153	53.745	7.95
SEDL	50	44.200	11.02	67	36.746	13.81
ALL SPONSORS	1031	45.999	12.65	1724	43.532	12.16

TABLE A-7

MEANS AND STANDARD DEVIATIONS OF FAMILY INCOME INDEX BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT				FT			
	$ Amount	N	$\bar{X}$	S.D.	$ Amount	N	$\bar{X}$	S.D.
FAR WEST LABS	5642.	175	11.886	7.74	4440.	209	8.134	6.30
UNIVERSITY OF ARIZONA	7270.	158	15.354	7.16	5131.	257	10.658	6.76
BANK STREET	6523.	63	13.540	6.69	5030.	114	10.307	6.38
UNIVERSITY OF OREGON	5721.	77	12.364	7.35	4228.	93	7.140	5.83
UNIVERSITY OF KANSAS	4487.	110	8.291	7.16	4354.	308	7.773	6.53
HIGH SCOPE FOUNDATION	5358.	142	11.317	7.15	4385.	167	7.928	6.58
UNIVERSITY OF FLORIDA	6262.	67	13.104	7.35	4507.	165	8.358	7.04
EDUCATIONAL DEVELOPMENT	6057.	95	12.716	7.00	5495,	142	11.592	7.02
UNIVERSITY OF PITTSBURGH	7112.	60	14.783	6.18	6501.	145	13.503	7.01
SEDL	5052.	50	10.260	7.15	4480.	67	8.269	6.94
ALL SPONSORS	5876.	997	12.353	7.45	4795.	1667	9.318	6.89

NOTE:

Code	Yearly Income	Code	Yearly Income	Code	Yearly Income	Code	Yearly Income	Code	Yearly Income
01	Under $1,500	06	$3,800 - 4,199	11	$5,200 - 5,699	16	$7,400 - 7,699	21	$9,200 - 9,799
02	$1,500 - 1,999	07	4,200 - 4,399	12	5,700 - 6,199	17	7,700 - 7,999	22	9,800 - 9,999
03	2,000 - 2,499	08	4,400 - 4,699	13	6,200 - 6,799	18	8,000 - 8,199	23	10,000 - 11,999
04	2,500 - 3,199	09	4,700 - 4,999	14	6,800 - 7,199	19	8,200 - 8,699	24	12,000 - 14,999
05	3,200 - 3,799	10	5,000 - 5,199	15	7,200 - 7,399	20	8,700 - 9,199	25	15,000 and over

The dollar amounts are interpolated values.

TABLE A-8

MEANS AND STANDARD DEVIATIONS OF MOTHER'S EDUCATION INDEX BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LAB	174	5.534	1.08	215	5.400	.96
UNIVERSITY OF ARIZONA	162	6.012	1.03	248	5.512	1.05
BANK STREET	66	4.576	1.72	108	5.259	.97
UNIVERSITY OF OREGON	78	5.679	1.08	91	5.330	1.03
UNIVERSITY OF KANSAS	100	5.270	1.04	287	5.314	1.06
HIGH SCOPE FOUNDATION	139	5.281	1.25	157	5.280	1.22
UNIVERSITY OF FLORIDA	71	5.324	1.01	163	5.301	1.02
EDUCATIONAL DEVELOPMENT	93	5.548	1.17	145	5.600	.89
UNIVERSITY OF PITTSBURGH	59	5.966	.90	152	5.901	.88
SEDL	49	4.755	1.45	65	4.769	1.40
ALL SPONSORS	991	5.471	1.22	1631	5.406	1.06

NOTE:

1 = no schooling
2 = kindergarten - 3rd grade
3 = 4th - 6th grade
4 = 7th - 8th grade
5 = high school, incomplete
6 = high school, complete
7 = some college, incomplete
8 = college, complete

Data source: Follow Through Parent Interview

TABLE A-9

FAMILY STRUCTURE
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	N	Mother & Father	Mother Only	Other	N	Mother & Father	Mother Only	Other
FAR WEST LABS	178	65.7%	32.0%	2.2%	219	54.3%	44.3%	1.4%
UNIVERSITY OF ARIZONA	166	75.9	21.7	2.4	260	66.5	30.0	3.5
BANK STREET	67	83.6	14.9	1.5	119	51.3	40.3	8.4
UNIVERSITY OF OREGON	82	58.5	37.8	3.7	95	35.8	60.0	4.2
UNIVERSITY OF KANSAS	113	60.2	28.3	11.5	320	47.5	45.6	6.9
HIGH SCOPE FOUNDATION	145	67.6	29.7	2.8	168	54.2	39.9	6.0
UNIVERSITY OF FLORIDA	74	78.4	17.6	4.1	173	55.5	39.9	4.6
EDUCATIONAL DEVELOPMENT	96	70.8	27.1	2.1	150	60.7	37.3	2.0
UNIVERSITY OF PITTSBURGH	60	93.3	5.0	1.7	153	90.2	9.8	0
SEDL	50	84.0	14.0	2.0	67	85.1	11.9	3.0
ALL SPONSORS	1031	71.5	25.0	3.5	1724	58.7	37.2	4.1

Data source: Follow Through Parent Interview

TABLE A-10

MEANS AND STANDARD DEVIATIONS OF NUMBER OF PERSONS IN HOUSEHOLD BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	178	5.854	2.36	219	5.740	2.22
UNIVERSITY OF ARIZONA	166	5.096	1.63	260	5.827	2.31
BANK STREET	67	5.254	1.58	119	5.571	2.04
UNIVERSITY OF OREGON	82	6.024	2.68	95	5.579	2.35
UNIVERSITY OF KANSAS	113	6.451	2.63	320	6.031	2.55
HIGH SCOPE FOUNDATION	144	5.354	1.99	168	6.077	2.39
UNIVERSITY OF FLORIDA	74	5.405	1.97	173	5.642	1.98
EDUCATIONAL DEVELOPMENT	96	6.094	2.29	150	6.087	2.45
UNIVERSITY OF PITTSBURGH	60	5.417	1.74	153	5.562	1.57
SEDL	50	6.320	2.06	67	6.866	2.50
ALL SPONSORS	1030	5.689	2.18	1724	5.868	2.28

Data source: Follow Through Parent Interview

Family income is a function of many things, not the least of which is the presence of a father and mother, both of whom may be wage earners. As can be seen in Table A-9, only 59 percent of the FT children come from homes in which both mother and father are present; this percentage is 72 among the NFT children in this sample. This leads us to suspect that the FT/NFT income differential is a function of the number of wage earners in the family, rather than of any intrinsic status differences between the families. The similarity in mothers' education tends to support this suspicion. Nevertheless, important psychodynamic differences must exist between families in which the father is missing and intact families.

Once again, there is wide variability both among and within Sponsors on this family structure variable. The University of Pittsburgh and SEDL programs work in communities where most of the families have both mother and father present. In this case of Pittsburgh this likely reflects the higher SES which characterizes this Sponsor's group. SEDL, on the other hand, includes very low income families so that the rather high rates of intact families probably reflects the ethnic values of predominantly Spanish-speaking communities. The University of Oregon's FT children come primarily from matriarchal homes although this is not true of the Oregon comparison children. The remaining Sponsors have slightly more of their FT children coming from intact families than from mother only homes. This differential is greatly magnified in the NFT group.

Table A-10 displays family size by Sponsor FT/NFT group. FT children overall come from significantly larger families than NFT children (FT=5.87, NFT=5.69). As with other SES indices, SEDL is on the extreme end of the Sponsor continuum with an average family size of 6.87. Two Sponsors--University of Arizona and High Scope--have significantly larger families in their FT than in their NFT groups.

While there is no significant difference overall between FT and NFT parents in their feelings of control over their destinies (see Table A-11 Parent General Locus of Control), three out of four significant differences within Sponsor favor the FT group. Parents in the Bank Street, University of Kansas, and University of Florida programs appear to feel more in control of their futures than parents in their respective NFT groups. On the other hand, parents in the University of Arizona program appear to feel less in control than their NFT counterparts.

TABLE A-11

MEANS AND STANDARD DEVIATIONS OF PARENT GENERAL LOCUS OF CONTROL BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	177	10.395	2.93	218	10.303	2.74
UNIVERSITY OF ARIZONA	165	11.133	2.50	259	10.525	2.83
BANK STREET	67	8.761	3.30	119	9.874	3.18
UNIVERSITY OF OREGON	82	9.732	3.01	95	9.568	3.06
UNIVERSITY OF KANSAS	113	8.947	2.97	320	9.787	2.96
HIGH SCOPE FOUNDATION	143	9.650	3.01	167	9.389	3.00
UNIVERSITY OF FLORIDA	73	9.027	3.50	172	10.349	2.95
EDUCATIONAL DEVELOPMENT	96	9.760	3.13	150	10.293	2.90
UNIVERSITY OF PITTSBURGH	60	10.417	3.04	153	10.980	2.66
SEDL	50	8.900	3.61	67	8.925	3.37
ALL SPONSORS	1026	9.863	3.11	1720	10.092	2.96

NOTE:
Scoring codes for items related to parent general locus of control:

Item	Response	Scoring Code
If parents do not agree with the people in our schools, there's not much parents can do.	Yes, agree No, disagree	2 5
No matter how hard a person tries, he can't do much about what happens to him.	Yes, agree No, disagree	1 4
A person shouldn't plan ahead because things don't usually work out.	Yes, agree No, disagree	1 4

Data source: Follow Through Parent Interview

From previous research, which is supported by correlations found in our NFT sample, we might expect a moderate positive relationship between income, education and parent Locus of Control. For Bank Street and the University of Arizona, attitudinal differences between the FT and NFT groups appear to be consistent with SES differences. Despite their lower incomes, the Bank Street mothers are better educated than their NFT counterparts and have more positive feelings about life in general. The University of Arizona parents, the most severely disadvantaged group relative to their NFT counterparts, also has far more negative attitudes. The positive attitudes of the University of Kansas and University of Florida FT parents, on the other hand, do not appear to be related to the SES characteristics of these groups. They may represent self-selection or early treatment effects. In either case, they suggest that for these Sponsors, the FT and NFT home environments may be qualitatively different from that found in families associated with other Sponsors.

Turning to parents' perceptions of the school's receptivity to their feelings and ideas (Table A-12), we find a significant difference overall in favor of the FT group. Two Sponsors contribute largely to this overall finding: University of Kansas and University of Florida. Once again, these differences could be attributed to either self-selection factors or early FT treatment effects.

4.0 TEACHER CHARACTERISTICS

Having described the pupils and their families, we now turn to a brief description of the teachers and classes in the analytic sample. We first discuss three characteristics of the classroom teacher--education, experience, and ethnicity--and later turn to the aggregated characteristics of the classroom children.

Table A-13 displays means and standard deviations for each Sponsor FT/NFT group on teacher education for the kindergarten year. Overall, there is a small, statistically significant difference between the FT/NFT groups on this variable, with FT teachers having received slightly less education than NFT teachers. There is some variability among Sponsor FT groups, with kindergarten teachers in the SEDL program being on the low end of the continuum, and teachers in the Bank Street, Kansas and EDC programs being on the high end. Within Sponsors there are also FT/NFT differences. FT teachers in the Arizona,

TABLE A-12

MEANS AND STANDARD DEVIATIONS OF PARENT PERCEPTION OF SCHOOL RECEPTIVITY BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	177	12.096	2.78	216	12.250	2.76
UNIVERSITY OF ARIZONA	165	12.127	2.91	255	12.212	2.72
BANK STREET	65	11.631	2.64	118	12.500	2.62
UNIVERSITY OF OREGON	81	11.395	2.61	95	11.968	2.71
UNIVERSITY OF KANSAS	113	11.027	3.10	316	12.025	2.76
HIGH SCOPE FOUNDATION	143	12.189	2.71	166	11.855	2.74
UNIVERSITY OF FLORIDA	72	11.333	3.34	170	12.253	2.73
EDUCATIONAL DEVELOPMENT	94	11.351	2.92	147	12.068	2.85
UNIVERSITY OF PITTSBURGH	60	11.683	3.49	152	12.559	2.78
SEDL	49	12.755	3.12	67	12.269	2.89
ALL SPONSORS	1019	11.795	2.95	1702	12.179	2.75

NOTE:
Scoring codes for items related to parent perception of school receptivity:

Item	Response	Scoring Code
The parents in your school have a say about how the schools are run.	Yes, agree No, disagree	5 2
The people in your school really know what the parents want.	Yes, agree No, disagree	6 3
The people in your school really care about what the parents think.	Yes, agree No, disagree	4 1

Data source: Follow Through Parent Interview

TABLE A-13

MEANS AND STANDARD DEVIATIONS OF LEVEL OF KINDERGARTEN TEACHER EDUCATION BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	178	3.938	.45	198	3.949	.52
UNIVERSITY OF ARIZONA	166	4.253	.76	260	3.554	.79
BANK STREET	67	2.935*	.86	114	4.132	.77
UNIVERSITY OF OREGON	82	4.585	.49	79	3.835	.54
UNIVERSITY OF KANSAS	98	3.918	.51	302	4.096	.72
HIGH SCOPE FOUNDATION	138	3.993	.66	153	4.013	.80
UNIVERSITY OF FLORIDA	74	3.676	.74	171	3.760	.59
EDUCATIONAL DEVELOPMENT	92	4.163	.66	143	4.161	.80
UNIVERSITY OF PITTSBURGH	54	3.537	.50	153	3.948	.62
SEDL	50	3.900	.30	56	3.339	.94
ALL SPONSORS	999	3.963	.71	1629	3.904	.74

NOTE:

1 = high school
2 = some college or technical training
3 = Bachelor's Degree
4 = some graduate level courses
5 = Master's Degree or beyond

*This value is correct; it appears that the the teachers in this site were certified before 4-year degrees were required in this state. The same is true at the first grade level in this site, but was obscured in the overall effect.

Data Source: Follow Through Teacher Questionnaire

Oregon, and SEDL programs are not as well educated as their NFT counterparts; FT teachers in the Bank Street, Kansas and Pittsburgh programs, on the other hand, are better educated than their NFT counterparts.

Turning to the first grade data (Table A-14), we find that, overall the education level of all teachers has decreased somewhat. It would appear that for a variety of reasons--including statewide certification policies, local screening and hiring practices, and self-selection--the kindergarten teachers in this subset are more apt to have earned advanced degrees than the first grade teachers.

Like the kindergarten teachers, the first grade FT teachers are slightly less educated than the NFT teachers. Within Sponsor differences also parallel those found in kindergarten. Teachers in the University of Kansas program are more highly educated than their NFT counterparts; FT teachers in the University of Arizona, University of Oregon, and SEDL programs are less well educated. While the differences are not statistically significant, Bank Street and Pittsburgh teachers are also somewhat more highly educated than their NFT counterparts at first grade level. Only EDC seems to have a slightly different pattern of teacher education across grades. In kindergarten, the FT/NFT teachers are equal in education; in first grade, the NFT group is superior.

Table A-15 displays means and standard deviations for teacher experience at the first grade level. Similar to the education pattern, we find that FT teachers overall are less experienced than the NFT teachers. Whereas the FT group as a whole has been teaching for approximately eight years, the NFT group has been teaching for more than twelve years. Variability among Sponsor FT groups is great, with teachers in the Oregon program being relatively new ($\overline{X}$=4.84) and those in the Far West and Pittsburgh programs relatively experienced ($\overline{X}$=10.63 and 10.05, respectively). With the exception of the Pittsburgh and SEDL programs, all Sponsor FT groups are less experienced than their respective NFT groups, and these differences are significant at the .05 level.

Education and experience have been used more frequently than any other criteria to screen and promote teachers. Salary schedules in most parts of the country are explicitly tied to them. If we use these traditional criteria to evaluate teacher quality, then we must conclude that the FT program is attracting and/or selecting less desirable teacher candidates. However, it

TABLE A-14

MEANS AND STANDARD DEVIATIONS OF LEVEL OF FIRST GRADE TEACHER EDUCATION
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.
FAR WEST LABS	150	3.387	.61	219	3.311	.53
UNIVERSITY OF ARIZONA	165	3.945	.85	239	3.808	.76
BANK STREET	67	3.716	.86	117	3.855	.87
UNIVERSITY OF OREGON	82	4.268	.58	95	3.874	.74
UNIVERSITY OF KANSAS	93	3.742	.64	302	4.089	.77
HIGH SCOPE FOUNDATION	129	3.605	.81	168	3.643	.56
UNIVERSITY OF FLORIDA	74	4.122	.72	171	3.556	.71
EDUCATIONAL DEVELOPMENT	83	4.024	.78	124	3.831	.96
UNIVERSITY OF PITTSBURGH	60	3.050	.50	146	3.130	.33
SEDL	29	3.517	.87	67	3.269	.86
ALL SPONSORS	932	3.750	.80	1648	3.677	.77

NOTE:
1 = high school
2 = some college or technical training
3 = Bachelor's Degree
4 = some graduate level courses
5 = Master's Degree or beyond

TABLE A-15

MEANS AND STANDARD DEVIATIONS OF FIRST GRADE TEACHER ELEMENTARY TEACHING EXPERIENCE IN YEARS BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	178	17.758	14.71	219	10.626	7.04
UNIVERSITY OF ARIZONA	166	12.645	11.37	260	6.081	6.64
BANK STREET	67	10.672	10.17	119	6.504	5.07
UNIVERSITY OF OREGON	80	11.837	5.85	83	4.843	3.34
UNIVERSITY OF KANSAS	98	13.735	10.69	320	8.912	8.71
HIGH SCOPE FOUNDATION	145	11.207	8.37	168	8.964	5.68
UNIVERSITY OF FLORIDA	74	13.635	8.43	173	9.324	8.83
EDUCATIONAL DEVELOPMENT	92	10.467	8.57	150	6.767	8.95
UNIVERSITY OF PITTSBURGH	60	10.700	8.98	153	10.052	11.05
SEDL	50	6.160	10.60	67	6.418	6.00
ALL SPONSORS	1010	12.688	10.94	1712	8.200	7.92

might be argued that less formally trained and experienced teachers may be more amenable to the demands of implementing Sponsors' models.

Turning to teacher ethnicity, we find that at the kindergarten level, there is no difference in the proportion of White teachers in the overall FT/NFT groups (FT=72.5%; NFT=75.1%). Table A-16 shows, however, a great deal of variability among Sponsor FT groups on this variable. Three Sponsors--University of Oregon, University of Kansas, and University of Florida--have roughly 50 percent of their FT teachers representing minority groups, two Sponsors--Bank Street and University of Pittsburgh--have fewer than 10 percent of their FT teachers representing minority groups. Furthermore, there are differences between the FT/NFT groups for several Sponsors. Bank Street and Kansas FT classrooms have more White teachers than their NFT counterparts. Far West Labs and University of Pittsburgh have approximately the same proportion; and the remaining Sponsors have somewhat fewer White teachers than their respective NFT groups.

The distribution of teacher ethnicity is very much the same in the first grade (Table A-17), as it is in kindergarten. There is an overall difference between FT/NFT groups, with the FT group having fewer White teachers (67.6%) than the NFT group (79.3%). Sponsor variability is also very similar to that found in the kindergarten group. Any difference between kindergarten and first grade pupil effects cannot easily be attributed to differences in teacher ethnicity between these grades for any Sponsor.

5.0 CLASSROOM CHARACTERISTICS

The major classroom characteristic discussed here is ethnic composition. We define a class with 90 percent or more minority children as predominantly minority, 90 percent of more White as predominantly White, and all classes in between as mixed.

The overall ethnic composition of FT classes in this sample is rather different in kindergarten than the NFT classes (Table A-18). There are more FT classes (47.1%) which are predominantly minority than NFT classes (24.7%). On the other hand, more NFT kindergarten classes (44.9%) are ethnically mixed than FT kindergarten classes (34.9%).

By first grade these cases are reduced somewhat (Table A-19). Of the FT first grade classes, 37.3 percent are predominantly minority and 29.1 percent of the NFT classes fall in this category. The two groups are equal in

TABLE A-16

KINDERGARTEN TEACHER ETHNICITY
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	Minority	White	N	Minority	White
FAR WEST LABS	178	21.3%	78.7%	198	23.7%	76.3%
UNIVERSITY OF ARIZONA	166	9.6	90.4	260	17.3	82.7
BANK STREET	67	26.9	73.1	114	9.6	90.4
UNIVERSITY OF OREGON	82	36.6	63.4	79	46.8	53.2
UNIVERSITY OF KANSAS	109	61.5	38.5	302	41.7	58.3
HIGH SCOPE FOUNDATION	138	25.4	74.6	153	32.0	68.0
UNIVERSITY OF FLORIDA	74	43.2	56.8	151	54.3	45.7
EDUCATIONAL DEVELOPMENT	92	10.9	89.1	119	21.0	79.0
UNIVERSITY OF PITTSBURGH	54	0	100.0	153	0	100.0
SEDL	50	10.0	90.0	56	25.0	75.0
ALL SPONSORS	1010	24.9	75.1	1585	27.5	72.5

TABLE A-17

FIRST GRADE TEACHER ETHNICITY
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	Minority	White	N	Minority	White
FAR WEST LABS	159	17.0%	83.0%	219	27.4%	72.6%
UNIVERSITY OF ARIZONA	165	4.8	95.2	239	27.6	72.4
BANK STREET	67	0	100.0	113	0.9	99.1
UNIVERSITY OF OREGON	82	56.1	43.9	95	53.7	46.3
UNIVERSITY OF KANSAS	89	37.1	62.9	291	48.8	51.2
HIGH SCOPE FOUNDATION	133	15.8	84.2	164	27.4	72.6
UNIVERSITY OF FLORIDA	74	40.5	59.5	171	52.0	48.0
EDUCATIONAL DEVELOPMENT	83	31.3	68.7	135	38.5	61.5
UNIVERSITY OF PITTSBURGH	60	0	100.0	146	0	100.0
SEDL	30	13.3	86.7	67	37.3	62.7
ALL SPONSORS	942	20.7	79.3	1640	32.4	67.6

TABLE A-18

KINDERGARTEN CLASSROOM ETHNICITY
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT				FT			
	N	Predominantly Minority	Mixed	Predominantly White	N	Predominantly Minority	Mixed	Predominantly White
FAR WEST LABS	172	27.3%	40.1%	32.6%	198	44.4%	48.5%	7.1%
UNIVERSITY OF ARIZONA	160	10.0	58.1	31.9	244	32.0	57.0	11.1
BANK STREET	52	11.5	53.8	34.6	100	21.0	18.0	61.0
UNIVERSITY OF OREGON	82	65.9	34.1	0	95	67.4	32.6	0
UNIVERSITY OF KANSAS	109	61.5	38.5	0	302	81.8	18.2	0
HIGH SCOPE FOUNDATION	138	21.7	67.4	10.9	153	51.6	48.4	0
UNIVERSITY OF FLORIDA	74	0	52.7	47.3	171	50.9	45.6	3.5
EDUCATIONAL DEVELOPMENT	92	31.5	3.3	65.2	143	64.3	14.7	21.0
UNIVERSITY OF PITTSBURGH	54	0	0	100.0	153	0	0	100.0
SEDL	50	8.0	92.0	0	56	7.1	92.9	0
ALL SPONSORS	983	25.7	44.9	29.4	1615	47.1	34.9	18.0

TABLE A-19

FIRST GRADE CLASSROOM ETHNICITY
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT				FT			
	N	Predominantly Minority	Mixed	Predominantly White	N	Predominantly Minority	Mixed	Predominantly White
FAR WEST LABS	145	35.2%	46.2%	18.6%	186	40.3%	54.8%	4.8%
UNIVERSITY OF ARIZONA	158	5.7	58.2	36.1	206	8.7	84.0	7.3
BANK STREET	52	57.7	0	42.3	87	9.2	10.3	80.5
UNIVERSITY OF OREGON	59	64.4	35.6	0	95	77.9	22.1	0
UNIVERSITY OF KANSAS	101	53.5	46.5	0	231	76.2	23.8	0
HIGH SCOPE FOUNDATION	131	35.9	60.3	3.8	154	42.2	57.8	0
UNIVERSITY OF FLORIDA	65	7.7	38.5	53.8	161	37.3	54.7	8.1
EDUCATIONAL DEVELOPMENT	66	22.7	0	77.3	97	54.6	14.4	30.9
UNIVERSITY OF PITTSBURGH	60	0	0	100.0	146	0	0	100.0
SEDL	29	10.3	89.7	0	67	6.0	94.0	0
ALL SPONSORS	866	29.1	41.2	29.7	1430	37.3	42.9	19.8

the percent of ethnically mixed (42.9% and 41.2%).

Within the FT group, there is a great deal of variability in class ethnic composition among Sponsors at both grade levels. Far West Labs, High Scope, and the University of Florida appear to work with both minority and mixed classes. Bank Street and the University of Pittsburgh appear to work primarily with White classes. The University of Oregon, University of Kansas and EDC work with minority classes. And finally, the University of Arizona and SEDL work primarily with mixed classes.

There are FT/NFT differences within Sponsor as well, and these differences vary by grade level. They are especially pronounced for Bank Street, whose White classes are matched primarily with mixed NFT classes in kindergarten and primarily with minority classes in first grade; University of Florida, whose minority and mixed classes are, for the most part, matched with White and mixed classes in both kindergarten and first grade; and EDC, whose minority classes are matched primarily with White classes at both grade levels. Although Sponsor differences and FT/NFT mismatches are discussed in terms of pretest differences, the ethnic differences across cells in the experimental design severely inhibit attributing Sponsor effects to their programs alone. The reader must keep this in mind when program effects are discussed below.

6.0 ENTERING PRETEST CHARACTERISTICS OF CLASSES

These entering levels[4] are assessed here by the Fall kindergarten WRAT (Table A-20). There is no overall difference between FT and NFT classes, a finding consistent with the pupil level entry data. FT/NFT classes within Sponsor do differ, however. As with the pupil level entry data, Far West Labs and University of Arizona classes start out with an initial disadvantage; University of Kansas classes with an initial advantage. In addition, the FT class means for both Bank Street and the University of Pittsburgh are above their NFT comparison class means in entering WRAT scores, although there was no corresponding advantage for these two Sponsors at the pupil level.

---

[4] Class entry level is the mean of all children in the class who exist in our traced longitudinal sample. This is a larger set than our ANCOVA analytic sample.

TABLE A-20

MEANS AND STANDARD DEVIATIONS OF CLASS MEAN FALL KINDERGARTEN WRAT BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	178	33.274	4.82	219	30.793	4.98
UNIVERSITY OF ARIZONA	166	35.870	7.86	260	31.261	5.03
BANK STREET	67	27.233	3.83	119	29.416	4.74
UNIVERSITY OF OREGON	82	30.828	6.61	95	32.114	4.29
UNIVERSITY OF KANSAS	113	26.431	4.26	320	29.669	4.30
HIGH SCOPE FOUNDATION	145	29.331	5.69	168	28.681	4.91
UNIVERSITY OF FLORIDA	74	29.258	3.60	173	30.467	5.17
EDUCATIONAL DEVELOPMENT	96	30.090	7.79	150	30.470	4.52
UNIVERSITY OF PITTSBURGH	60	35.316	3.57	153	37.028	6.03
SEDL	50	27.055	2.64	67	25.522	3.70
ALL SPONSORS	1031	31.033	6.54	1724	30.715	5.37

7.0 SCHOOL CHARACTERISTICS

When we turn to the school throughputs--or characteristics of the peers in the child's school--we find further FT/NFT differences. In terms of the ethnic composition of the schools in the analytic subset (Table A-21), the school level data (which come from the OCR directory) are quite similar to those found at the class level in kindergarten. Overall, the FT children in this subset are more apt to attend primarily minority schools than the NFT children and less apt to attend either integrated or primarily White schools. The pattern of school ethnic composition for FT Sponsor groups resembles that of class ethnic composition also, suggesting that for the most part, the classroom assignments of children within schools was not markedly altered in preparation for the FT experiment. For one Sponsor, however, the ethnic composition of the NFT classes differs dramatically from that of the NFT schools tested. Bank Street's NFT schools are more than 90 percent White; whereas at first grade, the NFT classes are over 57 percent minority.

Turning to school inputs, we find that the FT schools, which serve a lower income group than the NFT schools, receive more in the way of Federal dollars and auxiliary services. Table A-22, displays the proportion of FT/NFT schools receiving at least $50 per pupil from ESEA, Title I funds. Given the national per pupil Title I expenditure of approximately $200, this number implies that at least one-fourth of the pupils in the school are Title I recipients or that the school has a very large per pupil allocation. In either case, Title I funds of this magnitude might be expected to make a difference in school inputs. As can be seen, more FT school (67.7 percent) than NFT schools (40.1 percent) have this level of Title I funding. This differential is consistent for eight out of the ten Sponsors, with only the University of Kansas and SEDL having no difference between their FT/NFT groups. There is great variability among Sponsors, however. Far West Labs, University of Kansas and EDC are working in FT schools which all receive at least this minimum level of Title I funding; SEDL, on the other hand, has none of its schools receiving a noticeable amount of Title I funds.

Tables A-23 through A-26 present the person days of professional services available to FT and NFT children in these schools. These services

TABLE A-21

SCHOOL ETHNICITY
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT				FT			
	N	Predominantly Minority	Mixed	Predominantly White	N	Predominantly Minority	Mixed	Predominantly White
FAR WEST LABS	178	26.4%	48.9%	24.7%	219	38.8%	37.4%	23.7%
UNIVERSITY OF ARIZONA	166	9.6	42.2	48.2	260	21.2	57.3	21.5
BANK STREET	67	9.0	0	91.0	119	33.6	0	66.4
UNIVERSITY OF OREGON	82	42.7	57.3	0	95	55.8	36.8	7.4
UNIVERSITY OF KANSAS	113	66.4	33.6	0	320	82.8	17.2	0
HIGH SCOPE FOUNDATION	145	25.5	65.5	9.0	168	47.0	50.0	3.0
UNIVERSITY OF FLORIDA	74	0	52.7	47.3	173	32.9	43.9	23.1
EDUCATIONAL DEVELOPMENT	96	37.5	0	62.5	150	66.0	14.0	20.0
UNIVERSITY OF PITTSBURGH	60	0	0	100.0	153	0	0	100.0
SEDL	50	8.0	92.0	0	67	6.0	3.7	0
ALL SPONSORS	1031	24,8	40.9	34.2	1724	42.7	32.8	24.5

TABLE A-22

TITLE I FUNDS PER PUPIL IN SCHOOL
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	Less than $50	$50 or more	N	Less than $50	$50 or more
FAR WEST LABS	128	49.2%	50.8%	159	7.5%	92.5%
UNIVERSITY OF ARIZONA	124	91.1	8.9	205	57.1	42.9
BANK STREET	61	68.9	31.1	119	31.1	68.9
UNIVERSITY OF OREGON	82	51.2	48.8	71	33.8	66.2
UNIVERSITY OF KANSAS	113	0	100.0	302	3.0	97.0
HIGH SCOPE FOUNDATION	105	55.2	44.8	147	34.0	66.0
UNIVERSITY OF FLORIDA	69	100.0	0	173	39.9	60.1
EDUCATIONAL DEVELOPMENT	79	51.9	48.1	133	9.8	90.2
UNIVERSITY OF PITTSBURGH	55	100.0	0	153	62.7	37.3
SEDL	15	100.0	0	67	100.0	0
ALL SPONSORS	831	59.9	40.1	1529	32.3	67.7

TABLE A-23

MEANS AND STANDARD DEVIATIONS OF MAN DAYS OF PROFESSIONAL SERVICES PER PUPIL PER WEEK BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	128	.049	.03	159	.102	.05
UNIVERSITY OF ARIZONA	150	.068	.03	205	.112	.04
BANK STREET	61	.056	.02	119	.116	.04
UNIVERSITY OF OREGON	58	.112	.07	68	.076	.05
UNIVERSITY OF KANSAS	39	.047	[illegible]1	208	.059	.01
HIGH SCOPE FOUNDATION	97	.143	.15	163	.117	.05
UNIVERSITY OF FLORIDA	69	.046	.02	158	.114	.13
EDUCATIONAL DEVELOPMENT	74	.061	.03	144	.070	.03
UNIVERSITY OF PITTSBURGH	28	.018	.00	110	.054	.00
SEDL	15	.031	.00	67	.072	.01
ALL SPONSORS	719	.071	.07	1401	.095	.06

TABLE A-24

MEANS AND STANDARD DEVIATIONS OF MAN DAYS OF INSTRUCTIONAL SERVICES PER PUPIL PER WEEK BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	128	.016	.01	159	.043	.02
UNIVERSITY OF ARIZONA	150	.019	.01	205	.051	.02
BANK STREET	61	.028	.01	119	.043	.02
UNIVERSITY OF OREGON	58	.051	.02	68	.036	.01
UNIVERSITY OF KANSAS	39	.020	.00	208	.024	.00
HIGH SCOPE FOUNDATION	97	.050	.05	163	.055	.03
UNIVERSITY OF FLORIDA	69	.010	.00	158	.054	.05
EDUCATIONAL DEVELOPMENT	74	.022	.01	144	.021	.01
UNIVERSITY OF PITTSBURGH	28	.005	.00	110	.021	.00
SEDL	15	.017	.00	67	.035	.00
ALL SPONSORS	719	.025	.02	1401	.039	.03

TABLE A-25

MAN DAYS OF SOCIAL SERVICES PER PUPIL PER WEEK
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	Man-Days per Week <.005	Man-days per Week ≥.005	N	Man-Days per Week <.005	Man-Days per Week ≥.005
FAR WEST LABS	128	79.7%	20.3%	159	23.3%	76.7%
UNIVERSITY OF ARIZONA	150	65.3	34.7	205	37.6	62.4
BANK STREET	61	75.4	24.6	119	11.8	88.2
UNIVERSITY OF OREGON	58	22.4	77.6	68	73.5	26.5
UNIVERSITY OF KANSAS	39	28.2	71.8	208	39.9	60.1
HIGH SCOPE FOUNDATION	97	51.5	48.5	163	44.8	55.2
UNIVERSITY OF FLORIDA	69	100.0	0.0	158	56.3	43.7
EDUCATIONAL DEVELOPMENT	74	31.1	68.9	144	63.2	36.8
UNIVERSITY OF PITTSBURGH	28	100.0	0.0	110	23.6	76.4
SEDL	15	100.0	0.0	67	0.0	100.0
ALL SPONSORS	719	63.3	36.7	1401	38.5	61.5

TABLE A-26

MAN DAYS OF MEDICAL SERVICES PER PUPIL PER WEEK

BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	Man-Days per Week <.005	Man-Days per Week ≥.005	N	Man-Days per Week <.005	Man-Days per Week ≥.005
FAR WEST LABS	128	70.3%	29.7%	159	0.0%	100.0%
UNIVERSITY OF ARIZONA	150	21.3	78.7	205	0.0	100.0
BANK STREET	61	6.6	93.4	119	0.0	100.0
UNIVERSITY OF OREGON	58	69.0	31.0	68	19.1	80.9
UNIVERSITY OF KANSAS	39	48.7	51.3	208	15.4	84.6
HIGH SCOPE FOUNDATION	97	56.7	43.3	163	20.9	79.1
UNIVERSITY OF FLORIDA	69	49.3	50.7	158	8.9	91.1
EDUCATIONAL DEVELOPMENT	74	31.1	68.9	144	7.6	92.4
UNIVERSITY OF PITTSBURGH	28	100.0	0.0	110	51.8	48.2
SEDL	15	73.3	26.7	67	0.0	100.0
ALL SPONSORS	719	63.3	36.7	1401	11.5	88.5

may be purchased through Follow Through, Title I, or other Federal, state or local dollars. As can be seen in Table A-23, the FT children, overall receive significantly more person days of professional service per week (.095 per child) than the NFT children (.071 per child). Put another way, for every 100 FT children, approximately two professionals are available on a full time, equivalent basis to provide supplementary intructional, library/media, social, psychological and medical services; for every 100 NFT children, only 1.2 professionals are available on a similar basis.

As with many other characteristics used to describe the analytic subset, there is variability among Sponsors on this variable. University of Kansas and University of Pittsburgh have relatively few person days of service available per FT child (.059 and .054 respectively), whereas Florida has relatively many (.144). Eight of the ten Sponsors have more professional services available in their FT than in their NFT groups, the two exceptions being the University of Oregon and High Scope.

If we examine three components (namely instructional, social and medical) which contribute to the overall professional services, we find some differences both among and within Sponsor groups. Without giving a detailed summary of these variations, (see Tables A-24 through A-26), it is apparent that: (1) different Sponsors have differing patterns of supplementary services, (2) across all Sponsors, FT children consistently receive more medical services than their NFT counterparts, but not necessarily more instructional and social services, and (3) for University of Oregon, University of Kansas, High Scope, and EDC, the amount of support services sometimes favor FT, sometimes NFT.

Comparing the FT/NFT groups on two other input characteristics--total school enrollment (Table A-27) and student teacher ratio (Table A-28)--we find once again systematic differences overall, among Sponsors, and within Sponsor FT/NFT groups.

Overall, FT schools are larger than NFT schools. EDC serves the largest schools in the FT subset ($\bar{X}$=962 children), University of Pittsburgh the smallest ($\bar{X}$=323 pupils). University of Arizona Bank Street, University of Kansas, High Scope and EDC serve larger schools than their respective NFT groups, whereas Far West Labs, University of Florida, and University of Pittsburgh serve smaller ones.

TABLE A-27

MEANS AND STANDARD DEVIATIONS OF TOTAL ENROLLMENT IN SCHOOL BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	128	619.445	193.54	159	486.811	218.38
UNIVERSITY OF ARIZONA	150	442.247	217.99	205	506.937	213.38
BANK STREET	61	306.393	20.40	119	578.227	374.83
UNIVERSITY OF OREGON	82	550.707	93.58	91	555.681	257.41
UNIVERSITY OF KANSAS	113	432.531	268.33	320	834.797	476.99
HIGH SCOPE FOUNDATION	105	561.095	224.50	163	828.828	248.53
UNIVERSITY OF FLORIDA	69	611.072	210.02	173	537.659	158.81
EDUCATIONAL DEVELOPMENT	86	630.767	410.44	144	962.021	328.60
UNIVERSITY OF PITTSBURGH	55	617.945	172.85	153	323.131	136.49
SEDL	15	851.267	503.80	67	897.776	160.71
ALL SPONSORS	864	532.906	259.46	1594	655.001	362.11

TABLE A-28

MEANS AND STANDARD DEVIATIONS OF STUDENT-TEACHER RATIO IN SCHOOL BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT			FT		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
FAR WEST LABS	178	28.947	3.36	219	22.473	3.19
UNIVERSITY OF ARIZONA	166	24.916	3.57	260	22.335	1.78
BANK STREET	67	23.682	2.29	119	20.927	2.03
UNIVERSITY OF OREGON	82	24.316	1.16	95	21.509	4.50
UNIVERSITY OF KANSAS	113	24.157	3.76	320	23.393	2.45
HIGH SCOPE FOUNDATION	145	25.890	5.13	168	23.099	9.00
UNIVERSITY OF FLORIDA	74	26.559	2.20	173	27.534	4.04
EDUCATIONAL DEVELOPMENT	96	22.756	4.06	150	27.184	3.22
UNIVERSITY OF PITTSBURGH	60	27.452	3.60	153	25.158	2.57
SEDL	50	24.350	2.59	67	27.139	1.03
ALL SPONSORS	1031	25.275	4.03	1724	23.862	4.44

Overall, the FT schools have a smaller student/teacher ratio than the NFT schools (FT=23.86, NFT=25.88). Bank Street schools have a low ratio relative to other FT groups. On the other hand, the University of Florida, EDC and SEDL have a relatively high ratio, both relative to other FT groups and their NFT counterparts.

The application of each model in a school may be strongly influenced by the characteristics of the schools. Since we do not know which Sponsors may have its implementation enhanced by a large or small school, we cannot easily utilize these school differences across Sponsors as we interpret pupil effects. The differences represent sources of unknown error in the analysis and these must be addressed ultimately before a real understanding of Sponsor effects on children can be achieved.

8.0 COMMUNITY CHARACTERISTICS

Finally, having observed differences in pupil characteristics, family background, class, and school factors we turn to a brief look at the communities in which the Sponsor FT/NFT groups operate. Tables A-29 A-30, present a picture of the analytic subset by region and city size. Here again, we find wide variability.

The analytic subset is not evenly distributed by geographic region. While the subset is primarily Northern in composition, there are great differences among Sponsors in community location. Bank Street and EDC children are located entirely in the Northeast, University of Oregon and University of Pittsburgh, primarily in the North Central region, and SEDL's FT children, primarily in the South. The children in the remaining Sponsors' FT groups are somewhat more scattered geographically: University of Arizona and University of Oregon are primarily operating in both northeastern and North Central communities, Far West Labs in primarily North Central and Western communities, and High Scope in all regions of the country. For most

TABLE A-29

REGIONAL DISTRIBUTION OF CHILDREN BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT					FT				
	N	North East	North Central	South	West	N	North East	North Central	South	West
FAR WEST LABS	178	3.9%	47.2%	0.0%	48.9%	219	13.2%	42.5%	0.0%	44.3%
UNIVERSITY OF ARIZONA	166	25.3	62.0	12.7	0.0	260	43.1	45.8	11.2	0.0
BANK STREET	67	100.0	0.0	0.0	0.0	119	100.0	0.0	0.0	0.0
UNIVERSITY OF OREGON	82	22.0	78.0	0.0	0.0	95	8.4	91.6	0.0	0.0
UNIVERSITY OF KANSAS	113	10.6	65.5	23.9	0.0	320	47.8	35.0	17.2	0.0
HIGH SCOPE FOUNDATION	145	4.1	21.4	25.5	49.0	168	16.1	22.0	17.9	44.0
UNIVERSITY OF FLORIDA	74	0.0	47.3	52.7	0.0	173	11.0	0.0	54.9	34.1
EDUCATIONAL DEVELOPMENT	96	100.0	0.0	0.0	0.0	150	100.0	0.0	0.0	0.0
UNIVERSITY OF PITTSBURGH	60	45.0	55.0	0.0	0.0	153	28.1	71.9	0.0	0.0
SEDL	50	8.0	0.0	22.0	70.0	67	6.0	0.0	94.0	0.0
ALL SPONSORS	1031	27.1	41.1	13.1	18.7	1724	38.5	32.4	15.8	13.3

TABLE A-30

DISTRIBUTION OF CHILDREN BY CITY SIZE
BY SPONSOR BY FT/NFT FOR ANCOVA ANALYTIC SAMPLE

	NFT					FT				
	N	less than 10,000	10,000 to 50,000	50,000 to 200,000	over 200,000	N	less than 10,000	10,000 to 50,000	50,000 to 200,000	over 200,000
FAR WEST LABS	178	0.0%	0.0%	77.0%	23.0%	219	0.0%	0.0%	71.7%	28.3%
UNIVERSITY OF ARIZONA	166	0.0	28.3	24.1	47.6	260	0.0	33.1	18.8	48.1
BANK STREET	67	0.0	26.9	64.2	9.0	119	0.0	23.5	42.9	33.6
UNIVERSITY OF OREGON	82	0.0	0.0	100.0	0.0	95	0.0	0.0	100.0	0.0
UNIVERSITY OF KANSAS	113	33.6	0.0	10.6	55.8	320	17.2	0.0	14.4	68.4
HIGH SCOPE FOUNDATION	145	7.6	30.3	0.0	62.1	168	0.0	39.3	0.0	60.7
UNIVERSITY OF FLORIDA	74	47.3	0.0	0.0	52.7	173	0.0	34.1	0.0	65.9
EDUCATIONAL DEVELOPMENT	96	5.2	46.9	40.6	7.3	150	4.0	8.7	35.3	52.0
UNIVERSITY OF PITTSBURGH	60	55.0	45.0	0.0	0.0	153	71.9	28.1	0.0	0.0
SEDL	50	92.0	0.0	0.0	8.0	67	94.0	0.0	0.0	6.0
ALL SPONSORS	1031	16.3	17.6	34.2	31.9	1724	13.6	17.1	26.2	43.2

Sponsors, the FT and NFT groups are similar in geographic distribution, reflecting the original sampling design. For others, noticeably University of Florida and SEDL, the FT/NFT groups of our ANCOVA analytic sample are drawn from different regions of the country. Kansas has a much higher percentage of its NFT children in the North Central than it does FT children; note also that the Kansas NFT sample (N=113) is much smaller than the FT sample (N=320).

Over 43 percent of the FT children and 32 percent of the NFT children in the subset reside in communities of 200,000 persons or more. Once again, however, there are differences among Sponsor groups. The University of Kansas, High Scope, University of Florida and EDC have a majority of their FT children in these large metropolitan communities: Far West Labs and the University of Oregon have a majority of their FT children in middle-sized cities of 50,000 to 200,000; and University of Pittsburgh and SEDL have children from primarily rural communities with populations of 10,000 or less. The Arizona FT children are primarily from either large metropolitan cities (over 200,000) or small-sized cities (10,000 to 50,000). A smaller percentage is found in the middle-sized cities (50,000 to 200,000). Bank Street also has a sizable proportion of its FT children in all three non-rural settings. Although most Sponsors have similar FT and NFT distributions of children across city sizes, the University of Florida and EDC FT children appear to be in markedly different cities than their NFT counterparts.

APPENDIX B

METHODOLOGICAL ISSUES

APPENDIX B

METHODOLOGICAL ISSUES

1.0 INTRODUCTION

The research questions which motivate these analyses are framed in the language of the analysis of covariance. Typically, they take the form, "Within Sponsor X, what effect did Follow Through have upon pupils' scores on an outcome measure as administered in the spring of 1973, taking account of the fact that Follow Through and non-Follow Through groups may have differed systematically in initial achievement status?" Or, equivalently, "Controlling for initial achievement, what Follow Through effects within Sponsor can we document?" In this illustrative sample, the score of an individual pupil on an outcome measure is used as a criterion variable; his initial achievement level is the covariate of the criterion variable; and the administration (or lack) of the Follow Through program is the "treatment". In this particular case, the interest centers around estimating the effect on the criterion variable of administering the treatment. Ideally, one would select a set of "otherwise equal" pupils; administer Follow Through to randomly chosen subsets of these; not administer Follow Through to the remaining subsets; and then measure the differences in some prespecified score, thus arriving at a direct measure of the Follow Through effect. This is the ideal case of an experimental design. In reality, it is practically impossible to define a set of "otherwise equal" pupils: There are differences in the socio-economic status of their parents; their ethnicity; learning potential; pre-formal school knowledge; and the prejudices gained through relatives, friends, neighbors, etc. The analysis of covariance is a statistical technique designed to compensate (at least partially) for the effects of these differences on the observed values of the criterion variable and thus to isolate the effects of deliberate design variations. Before describing the analysis of covariance as a technique for evaluating the Follow Through effects, we will summarize the underlying multiple linear relationship model and its implications.

2.0 THE GENERAL LINEAR RELATIONSHIP MODEL

In mathematical terms, the covariance technique is a particular case of the general linear model wherein the covariates, treatments, etc., together form the set of "predictors", alternately called "right-hand-side variables", "regressors", and "independent variables". The values of the criterion variable (i.e., the dependent variable) are expressed as a linear function of these predictors, together with a "normally distributed" error term. A brief review of the linear model, therefore, precedes the discussion of the analysis of covariance.

2.1 Linearity Assumptions

Formally, the general linear model is described by the joint set of equations:

$$(1) \qquad Y_j = B_0 + B_1X_{1j} + \ldots + B_kX_{kj} + E_j \qquad (j = 1, 2, \ldots, n)$$

where Y_j is the j^{th} observation on the dependent variable Y; X_{ij} $(i = 1, 2, \ldots, k)$ is the concomittant observation on the i^{th} predictor, variable X_i; and E_j is the error of prediction in the j^{th} observation on Y. The classical statistical theory assumes that these errors are normally and independently distributed with mean zero and variance σ_E^2.

This innocuous-looking statement embodies the main assumptions underlying the use of a linear model, some of which are discussed below.[1]

1. The dependent variable, Y, is continuously measurable even though, by necessity, the observed values of Y are discontinuous.
2. Y can assume values over the entire range of real numbers, i.e., from negative infinity to positive infinity. While studying social systems, this assumption is often violated (Y varies over some limited range), apparently without much serious consequence. In the most severe instances, e.g., when Y denotes probability of success (and thus cannot assume values outside of the [0, 1] interval), certain "tail-stretching" transformations are customarily employed (such as arctan, probit, logit).

---

[1] These assumptions are not unknown. It is, however, very easy to be intrigued by the fancy formulae and lose sight of the underlying assumptions.

3. The effects of the right-hand-side variables are additive, as implied by the functional form of equation (1). In some special cases, this assumption is met by transforming the variables defining equation (1). For example, in some biological processes, the effects are multiplicative rather than additive—population growth being a prime example. The additivity of the effects is achieved by employing a logarithmic transformation. For example, the relationship $Y_j = B_0 \cdot X_{1j}^{B_1} \cdot X_{2j}^{B_2} \cdot \ldots \cdot X_{Kj}^{B_k} \cdot f_j$ is expressed as: $\log Y_j = \log B_0 + B_1 \log X_{1j} + \ldots + B_k \log X_{kj} + E_j$. Similarly, the relationship: $Y_j = B_0 + B_1 X_j + \ldots + B_k X_j^k + E_j$ is expressed by the equation: $Y_j = B_0 + B_1 X_{1j} + \ldots + B_k X_{kj} + E_j$ where $X_{ij} = X_j^i$.

   It must be emphasized that the additive effects are postulated only on these transformed variables, and not on the initial variable set.
4. The "normality" assumption is the most difficult assumption to either justify on theoretical grounds or verify in practical situations; general recourse is to assume the appropriateness of the "central limit theorem", particularly when the Y values represent some kind of averaging.
5. The assumption that the "error terms are independent of each other" is tantamount to stating that the observations are statistically pure. The theory of statistical inferences and the possibility of generalizing the research findings rest squarely on this assumption. If this assumption fails, no transformation or other statistical gimmicks can alter the fact of nonrandomness.

2.2 Correlations: Partial, Multiple, and Semi-partial

Provided that those statistical assumptions are met, the classical multivariate analysis techniques are applicable. Summarized below are some of the main results and formulae which are used while evaluating the Follow Through effects.

1. Let r_{uv} denote the zero-order correlation between two variables u and v. Then the partial correlation $r_{uv \cdot w}$, i.e., the correlation between u and v when the variations due to w have been accounted for, is computed by the recursive formula:

$$r_{uv.w} = \frac{r_{uv} - r_{uw} \cdot r_{wv}}{\sqrt{1 - r_{uw}^2} \cdot \sqrt{1 - r_{wv}^2}} \tag{2}$$

Consequently, if any two of these three variables are perfectly correlated, all but one of the resultant partial correlations are meaningless. For example, $r_{uv} = 1$ implies that $r_{uw} = r_{wv}$ so that $r_{uv.w} = 1$. On the other hand, the denominators appearing in the formulae for computing $r_{uw.v}$ or $r_{vw.u}$ are zero. Similar remarks are true when $r_{uv} = -1$.

The recursive formula (2) generalizes when there are four or more variables. For example,

$$r_{uv.wz} = \frac{r_{uv.w} - r_{uz.w} \cdot r_{vz.w}}{\sqrt{1 - r_{uz.w}^2} \cdot \sqrt{1 - r_{vz.w}^2}} \tag{3}$$

The comments regarding perfect correlations continue to hold with an added feature: Perfect correlations of first order are generally harder to anticipate than those of the zero order. Consequently, while trying to "fit Y on $X_1, \ldots, X_k$" by employing computer-based statistical packages, the analyst is often surprised to see computer messages of the type, "Partial correlation equals unity, cannot proceed." Whenever this happens, all but one of the perfectly correlated predictors must be deleted from further analyses.

2. The maximum correlation between the dependent variable, Y, and the linear construct of predictors, $B_0 + B_1 X_1 + \ldots + B_k X_k$, is the multiple correlation coefficient, $R_{y.12 \ldots k}$ where:

$$R^2_{y.12 \ldots k} = \begin{bmatrix} r_{y1} & \cdots & r_{yk} \end{bmatrix} \begin{bmatrix} 1 & r_{12} & \cdots & r_{1k} \\ \vdots & & & \\ r_{21} & 1 & \cdots & r_{2k} \\ \vdots & & & \\ r_{k1} & r_{k2} & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_{y1} \\ \vdots \\ r_{yk} \end{bmatrix} \tag{4}$$

If, for some pair of distinct indices i1 and i2, $r_{i1,i2} = 1$, then the central matrix on the right-hand side of Equation 4 is singular, R^2 of some predictor on other predictors is 1, and $R^2_{Y\cdot12\ldots k}$ is meaningless. Similar remarks hold if the partial correlation coefficient between a pair of predictors equals one. Conversely, if R^2 of some predictor on other predictors equals one, then at least one of the partial correlations or zero-order correlations is likely to be unity, thus resulting in a meaningless $R^2_{Y.12\ldots k}$.

It can be shown that, when the population value of R^2 is zero, its sample counterpart, $\hat{R}^2$, is such that $E(\hat{R}^2) = \frac{k}{n-1}$ and that $\mathrm{Var}(\hat{R}^2) = \frac{2(n-k)(k-1)}{(n^2-1)(n-1)}$.

Thus, when the number of predictors, k, approaches the number of cases, n, $\hat{R}^2$ may become spuriously large, and the conclusions drawn from the sample can be rather unstable. This difficulty is minimized if k is sufficiently small relative to n. The threshold value for the upper limit on the spurious $\hat{R}^2$ is subjective; k/n = .05 may be safe, k/n = .10 might be more achievable in some studies. Yet another alternative, for generalization of results after drawing statistical inferences on the basis of the sample, is to employ the more conservative estimate of $R^2_{Y.12\ldots k}$ (the "shrunk" R^2, symbolically sR^2). This is computed by the formula:

$$sR^2_{Y.12\ldots k} = 1 - (1 - \hat{R}^2_{Y.12\ldots k})\,\frac{n-1}{n-k-1}\,. \quad (5)$$

That the equation 5 provides adequate protection against spurious correlations may be seen as follows. When true $R^2_{Y.12\ldots k}$ equals zero, the expected value of its sample estimate, $E(\hat{R}^2_{Y.12\ldots k})$, is given by:

$$E(\hat{R}^2_{Y.12\ldots k}) = \frac{k}{n-1}\,. \quad (6)$$

It follows from (5) and (6) that when $R^2_{Y.12...k} = 0$, $E(sR^2_{Y.12...k}) = 0$.

Thus, *after drawing statistical inferences* about the strength of relationship, it is *always* preferable to use $sR^2_{Y.12...k}$ instead of $\hat{R}^2_{Y.12...k}$ for purposes of generalizing the findings.

3. The B-weights, B_0, B_1, ..., B_k are estimated by the formula:

$$\hat{B}_i = \hat{r}_{yi.c} \frac{s_{y.c}}{s_{i.c}} \tag{7}$$

where c denotes that subset of subscripts 1, 2, ..., k from which i has been excluded; $s_{y.c}$ and $s_{i.c}$ are partial standard deviations of Y and X_i respectively. Furthermore, $s^2_{i.c} = s^2_i\ (1 - R^2_{i.c})$ and $s^2_{y.c} = s^2_y\ (1 - R^2_{Y.c})$ are partial variances. Finally, the estimate of the additive constant, $\hat{B}_0$, equals:

$$\hat{B}_0 = \bar{Y} - \hat{B}_1\bar{X}_1 - \ldots - \hat{B}_k\bar{X}_k.$$

4. The null hypothesis

$$H_0\colon B_1 = B_2 = \ldots = B_k = 0$$

is equivalent to the null hypothesis that Y values are unrelated to the predictors X_1, ..., X_k, i.e., the null hypothesis:

$$H_0\colon R^2_{Y.12...k} = 0.$$

The latter hypothesis is tested by the F-ratio:

$$F = \frac{n-k-1}{k} \frac{\hat{R}^2_{Y.12...k}}{1 - \hat{R}^2_{Y.12...k}} \tag{8}$$

which has k and n-k-1 degrees of freedom.

5. The hypothesis that the i^{th} B-weight, B_i, has a *prespecified* value B^0_i is tested by the t-ratio:

$$t = \sqrt{n-k-1}\ \cdot\ \frac{s_{i.c}}{s_{y.i,c}} \left[B_i - B^0_i \right] \tag{9}$$

In this formula, $s^2_{i.c} = s^2_i \, (1 - R^2_{i.c})$ and $s^2_{y.i,c} = s^2_{y.12...k} = s^2_y \, (1 - \hat{R}^2_{Y.12...k})$. The important point to remember is that the hypothetical value $\underline{B}^0_i$, must always be chosen for reasons external to the sample at hand or prior to an inspection of its estimated value.

2.3 Three Variations of the Linear Model

There are three distinct variations of the general linear model: the analysis of regression, the analysis of variance, and the analysis of covariance. The first two are discussed below, and the analysis of covariance is discussed in a separate section.

1. Analysis of Regression. The interest lies in predicting the values of the dependent variable Y as a function of the independent variable set $X_1, X_2, \ldots, X_k$. In this situation, the linear model is referred to by the name "linear regression model", and the B-weights are called the "regression coefficients." Traditionally, the regression model has assumed that not only Y but $X_1, \ldots, X_k$ are also continuous, and that in fact, $(Y, X_1, \ldots, X_k)$ follow a multivariate normal distribution. This model has not been used while evaluating the Follow Through effects; instead, we have coded the predictors in a manner analogous to the fixed effects analysis of variance.

2. Analysis of Variance. The interest lies in predicting Y as a linear function of $X_1, \ldots, X_k$, but the predictor variables are finite-valued and represent experimental treatments administered to "otherwise equal" subjects, i.e., deliberate design variations. Some of the predictors represent the main factors, others are called the interactions among the main factors. In the 2^m factorial design, the predictors are usually chosen so that their coefficients constitute an orthogonal set, the number of predictors being equal to the total degrees of freedom in the associated experimental design. For example, a 2 x 2 factorial design

incorporating two levels (A_1, A_2) of factor A and two levels (B_1, B_2) of factor B has three degrees of freedom, one for each of the main factors and one for the interaction. The treatment combinations and the corresponding "effects coding scheme" is represented below.

Treatments	Predictors			Frequency
	X_1	X_2	X_3	
A_1 B_1	1/2	1/2	1/4	n_{11}
A_1 B_2	1/2	-1/2	-1/4	n_{12}
A_2 B_1	-1/2	1/2	-1/4	n_{21}
A_2 B_2	-1/2	-1/2	1/4	n_{22}

In this scheme, X_1 represents treatment A effects; X_2 represents treatment B effects and X_3 represents interaction A x B effects.

If $n_{11} = n_{12} = n_{21} = n_{22}$, or if $\frac{n_{11}}{n_{11} + n_{12}} = \frac{n_{21}}{n_{21} + n_{22}}$, then (and only then) X_1, X_2, and X_3 are uncorrelated, and thus the factor effects are not confounded. In the traditional analysis of variance schemes, confounding of the factors is a nuisance; the multiple regression framework employed while evaluating the Follow Through effects provides an efficient method of dealing with correlated factors.

Since the "regression coefficients", B_1, B_2, and B_3 equal the factor A effect, factor B effect, and the interaction A x B effect, respectively, corresponding analysis of variance null hypotheses are equivalent to $B_1 = 0$, $B_2 = 0$ and $B_3 = 0$, respectively. These latter hypotheses are tested as follows. Let $sr_i^2 = R^2_{Y.i,c} - R^2_{Y.c}$, where c is that subset of subscripts 1, 2, ..., k from which i has been excluded. Then sr_i^2 represents the unique contribution of X_i towards the reduction in the Y variance. If the null hypothesis, $B_i = 0$, is true, then the F-ratio:

(10) $$F_{1,n-k-1} = (n-k-1) \cdot \frac{sr_i^2}{1 - \hat{R}^2_{y.1...k}} ,$$

has Snedecor's F-distribution with 1 and n-k-1 degrees of freedom and is free of the confounding effects. An additional advantage of this coding scheme is that it encourages us to draw statistical inferences about the "overall strength of relationship," the sample estimate of this parameter being

(11) $$sR^2_{Y.12...k} = 1 - (1 - \hat{R}^2_{y.12...k}) \frac{n-1}{n-k-1} .$$

Equation (11) is seldom, if ever, used in the traditional analysis of variance.

3.0 THE ANALYSIS OF COVARIANCE

This technique is similar to the analysis of variance in that some of the predictor variables $X_1, \ldots, X_k$ represent main factors and interactions, and it is a particular case of the general linear model in that the interest lies in predicting Y as a linear function of these predictors. However, in the analysis of covariance, we do not assume the experimental subjects to be "otherwise equal". For example, the effects of administering Follow Through are measured by the differences in the values of the criterion variables such as the pupil's score on Spring WRAT, MAT, Gumpgookies, Locus of Control, and Absence Rate. But the pupils participating in the study were not chosen to be otherwise equal, nor can this be practically achieved "after the fact". Consequently, the observed differences among the pupils' scores are attributable not only to the differences in treatments and their interactions, but also to the intrinsic differences among pupils: There are differences in factors such as socio-economic status of their parents; their learning potential; pre-formal school knowledge (as measured, for example, by their scores on Fall WRAT); and the prejudices gained through relatives, friends, neighbors, etc. (as measured partly by variables such as the size of their residential town and geographical region of the country). The analysis of covariance is a statistical technique for isolating the treatment effects from the effects of these intrinsic differences. The analysis of covariance is thus the most suitable method which may be used to evaluate the Follow Through effects.

3.1 The Noise and the Signals

The basic linear model, as represented by equation 1, treats all of the predictors equally. As long as the predictor set $X_1, \ldots, X_k$ contains non-redundant predictors only (that is, so long as no pair of predictors is perfectly linearly correlated), the proportion of Y-variance accounted for by the linear model is $R^2_{y.12\ldots k}$ and the proportion of Y-variance not explained by the model is $1 - R^2_{y.12\ldots k}$. The latter term is interpreted as the "noise" in the model. Since the k + 1 parameters, B_0, B_1, ..., B_k, are estimated on the basis of n observations · $(Y_1, X_{11}, \ldots, X_{k1})$, $(Y_2, X_{12}, \ldots, X_{k2})$, ... $(Y_n, X_{1n}, \ldots, X_{kn})$, the noise in the data is generated by (n - k - 1) degrees of freedom; these degrees of freedom are used to estimate the error variance, σ^2_E. Consequently, the ratio $(1 - \hat{R}^2_{y.12\ldots k})/(n - k - 1)$ can be interpreted as the noise in the data per error degree of freedom. This ratio is the one invariant in the linear model: regardless of any interpretation imposed on any individual predictor or on a group of predictors, the ratio $(1 - \hat{R}^2_{y.12\ldots k})/(n - k - 1)$ is always an unbiased estimate of a multiple of σ^2_E. Any "signal," generated by one or more predictors is, therefore, always compared with this "unit noise".

The "total signal" coming out of the model is $R^2_{y.12\ldots k}$. Since this total is based on the k predictors comprising the model, the ratio: $\hat{R}^2_{y.12\ldots k} / k$ is interpreted as the average signal per predictor in the data. The model is said to be a "significant" explanation of the Y-variance if the signal-to-noise ratio in the data exceeds a predefined constant, i.e., if:

$$(12) \quad F = \frac{\text{Average Signal}}{\text{Unit Noise}} = \frac{n - k - 1}{k} \cdot \frac{\hat{R}^2_{y.12\ldots k}}{1 - \hat{R}^2_{y.12\ldots k}} > F_0 \ .$$

The Critical Level of Signal. Traditionally, the predefined value of the constant, F_0, is obtained from the statistical tables of Snedecor's F distribution with k and n - k - 1 degrees of freedom and is such that:

$$(13) \quad \text{Prob.} \left\{ F > F_0 \text{ when } R^2_{y.12\ldots k} = 0 \right\} = \alpha \ .$$

The "significance level", α, is generally chosen to be either .05 or .01. It must be emphasized that, apart from the applied sciences' tradition, there is no particular reason to choose $\alpha = .05$ or $\alpha = .01$. Ideally, the "critical" value, F_0, should be set only after considering both the sensitivity and the selectivity of the particular model. In other words, one should resort to the decision-theoretic approach for assessing the importance of the observed signal-to-noise ratio. Even the weakest and operationally meaningless signal will rise above the unit noise if the linear model is based on a sufficiently large number of observations. If the inequality as defined by (12) is satisfied, the only reasonable conclusion in such a case would be that "something is there." One must then compute $sR^2_{y.12...k}$ of equation 5 to determine the strength of the signal and thus the usefulness of the model. This issue becomes very critical while deciding on the heterogeneity of the covariates, and we will continue this discussion in that section.

Individual Signals. The semi-partial correlation, sr_i, is defined as the correlation of X_i from which all other predictors have been partialled out (literally, $X_{i.c}$) with an unpartialled Y. In other words, sr_i^2 is the amount by which $\hat{R}^2_{y.12...k}$ would be reduced in the data if X_i were not in the linear model:

$$(14) \quad sr_i^2 = \hat{R}^2_{y.i,c} - \hat{R}^2_{y.c} \; .$$

Analogous to the definition of total signal, sr_i^2 can be interpreted as the signal associated with the predictor X_i; this signal does not depend on the values of the other predictors, and thus is the unique signal associated with X_i given the presence of the other predictors. (Note that the total signal, $R^2_{y.12...k}$, is not the sum of these individual signals, $sr_1^2 + \ldots + sr_k^2$, unless the predictors, $X_1, \ldots, X_k$, are uncorrelated.) The signal-to-noise ratio defined in equation 10 can be used to test the strength of this signal. If this ratio is found to be significant, one can use the difference: $sR^2_{Y\cdot i,c} - sR^2_{Y\cdot c}$ for generalizing the results.

3.2 The Covariate Set

Yet another interpretation of sr_i^2 is possible. Consider the linear model in which there is *only one treatment*, and where this treatment is administered at one of the two levels (for example, Follow Through and non-Follow Through). Let X_i denote the effects-coded representation of this treatment. *Then sr_i^2 is the signal associated with this treatment when all other explanatory variables have been partialled out of X_i*. The set c, in this case, denotes the set of indices representing *other* variables which also contribute to the explanation of Y-variance but which are not deliberate design variations. In other words, the set of variables, $X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k$ together form a *covariate set*, and sr_i^2 is the effect of the treatment after adjusting for other (and unintended) variations. The basic strategy in the analysis of covariance is to compare this signal with the unit noise as defined earlier. The associated signal-to-noise ratio is given by equation 10, and is rephrased below.

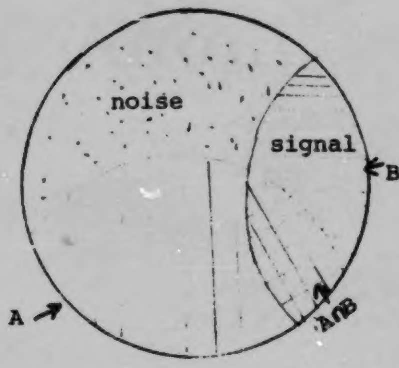
$$(15) \qquad F = \frac{(n-k-1)\,sr_i^2}{1 - \hat{R}^2_{y.i,c}} = \frac{n-k-1}{1} \cdot \frac{\hat{R}^2_{y.i,c} - \hat{R}^2_{y.c}}{1 - \hat{R}^2_{y.i,c}}$$

This strategy of accounting for "other variations" generalizes when there is more than one factor and/or when the factors appear at more than two levels.

Suppose that the predictor set $X_1, \ldots, X_k$ is divided into two mutually exclusive and exhaustive subsets A and B. Then, depending on one's purpose and the coding scheme, either A or B may represent the "treatment" set and the other will be called the covariate set. Equation 15 generalizes as follows. Let $\hat{R}^2_{y.AB}$ denote the total signal when both sets A and B are in the linear model and let $\hat{R}^2_{y.A}$ denote the signal when only the set A is in the model. Then the effect of B, after adjusting for the effects of A, is tested by the ratio:

$$F = \frac{\hat{R}^2_{y.AB} - \hat{R}^2_{y.A}}{1 - \hat{R}^2_{y.AB}} \cdot \frac{n-a-b-1}{b} \qquad (16)$$

where a = number of predictors in set A and b = number of predictors in set B. The sets A and B in equation 16 can be exchanged if their roles are reversed. Using the Venn-diagram of Y-variance, one may represent the signal-to-noise ratio as follows. The totality of the Y-variance is represented by the circle. The proportion, $\hat{R}^2_{y.B}$, of this variance is explained by the treatment set B (represented by the horizontal lines). This explanation, however, is contaminated by the joint effects of A and B, as indicated by the cross-hatched portion of that figure. Together, the treatment set (B) and the covariate set (A), explain a proportion, $\hat{R}^2_{y.AB}$, of the Y-variance--the most that can be explained with only A and B predictor sets in the model. Thus, the dotted area represents total noise, $1 - \hat{R}^2_{y.AB}$, and the horizontally shaded (but not cross-hatched) area represents the uncontaminated signal associated with set B. When adjusted for the appropriate degrees of freedom, the F of equation 16 represents the ratio of this signal to noise.



Since set A and B are not orthogonal as indicated by the intersection of A and B, the signal associated with B may in fact underestimate the true B effect. However, it is not possible to attribute the variance accounted for by A∩B to either A or B. This indeterminancy poses a logical difficulty but despite this problem, the F ratio of the unique effect of B is still an appropriate statistical test.

3.3 The Desirability of a Covariate Set

The reason for employing covariate(s) while assessing the effects of a treatment is that the covariates act as "statistical equalizers," and thus allow us to estimate the true effect of a treatment in spite of the necessarily unavoidable pre-treatment differences between the subjects receiving different levels of that treatment. Thus, to be useful in this regard, a covariate (or each covariate in a set) should possess the following properties:

1. It should account for an operationally meaningful[2] proportion of the criterion variance. We have refrained from using the language associated with probabilistic sampling. For example, we do not wish to say that "the explanation offered by the covariate (set) should be statistically significant," although the computations are similar. The language incorporating words such as "significant" is also avoided because it tends to give the illusion that there is something sacrosanct about $\alpha = .05$ or $\alpha = .01$.
2. It should be homogeneous with respect to treatment(s). It should treat each level of (each) treatment alike, i.e., there should not be excessive interaction between a covariate and treatment(s). This is important, since an interaction implies that the mean of the observed Y-values depends jointly on the covariate level and the treatment level, i.e., the covariate is not acting as an equalizer.

For each potential covariate, the first condition can be tested by computing the signal-to-noise ratio in the data:

$$F = \frac{\hat{R}^2_{y.A}}{1 - \hat{R}^2_{y.A}} \cdot \frac{n-a-1}{a}, \tag{17}$$

and then deciding whether it is sufficiently large.

---

[2] Due to the lack of a well-developed decision theoretic model, we will continue to use $F_{nd,dd;\alpha}$ as the surrogate for the threshold value of the signal-to-noise ratio (nd = numerator degrees of freedom, dd = denominator degrees of freedom and $\alpha = .05$, or .01). The need for such a model is ever more apparent.

If a covariate set does not contribute to the explanation of the criteria-variance, it is hardly worth keeping in the later analyses.

The second requirement can be tested by the ratio:

$$(18)\quad F = \frac{\hat{R}^2_{y\cdot A,B,AxB} - \hat{R}^2_{y\cdot A,B}}{1 - \hat{R}^2_{y\cdot A,B,AxB}} \cdot \frac{n - a - b - ab - 1}{ab}$$

It is more difficult to determine the proper action to be taken, if a covariate is inhomogeneous (i.e., if the ratio in (18) is excessive).

1. One may decide not to use such a covariate since its action is non-uniform across treatment levels, i.e., it is not doing its job. This is perhaps the safest, though not necessarily the most useful, solution. The covariate in question might have been extremely useful for explaining the Y-variance, and deleting it from further analyses may be a serious loss.

2. One may experiment with several non-linear transformations of the offending covariate in the hopes of finding one that 1) continues to correlate with the criterion variable, and 2) is less heterogeneous than the current covariate. The analyst's judgement and good fortune play an important role.

3. Use the fact that the covariate set is defined by exclusion ("that which is not a treatment set"). For example, let A denote the offending covariate, B denote the treatment set of the analyst's interest and let A x B denote the covariate x treatment interaction set. Then the joint set (A, A x B) can be viewed as the new covariate set. Consequently, the treatment effects can be tested by the ratio:

$$(19)\quad F = \frac{\hat{R}^2_{y.A,AxB,B} - \hat{R}^2_{y.A,AxB}}{1 - \hat{R}^2_{y.A,AxB,B}} \cdot \frac{n-a-b-ab-1}{b}\ .$$

Even though statistically appropriate, this procedure has many impractical features. These are 1) reduction in the error degrees of freedom, and corresponding danger of increased instability (see, for example, equation 8); 2) increase in computer operations costs--if the covariate is represented by $\underline{a}$ predictors and the treatment set by $\underline{b}$ predictors, the new predictor set (jointly) contains $(a + b) + a \cdot b$ predictors, and corresponding correlation matrix contains at least $2(a + b)ab$ extra terms; 3) increase in the lack of precision of the computer routine; and 4) increase in the difficulty of interpreting results. Furthermore, there is no assurance that the new covariate, viz. A x B, will not be found heterogeneous.

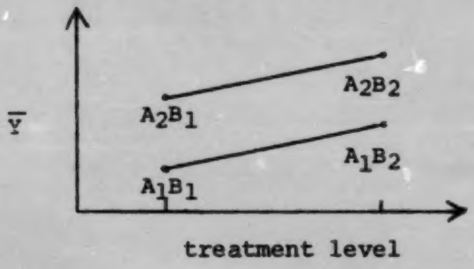
3.4 Comparison of Means

An alternative way of dealing with the covariate x treatment interaction is to examine the appropriate group means, and thus document the judgmental issues. If the covariate, A, is represented by a two-valued predictor and the treatment, B, is also represented by a two-valued predictor, there are four groups of interest, symbolically: A_1B_1, A_1B_2, A_2B_1, and A_2B_2. The observed interaction is computed to be:

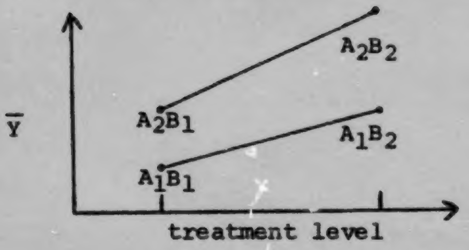
$$\widehat{A \times B} = \sum_{i,j} (\overline{Y}_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..})^2, \tag{20}$$

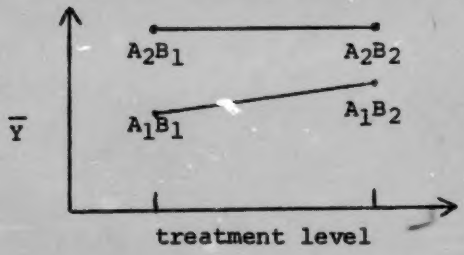
where $\overline{Y}_{ij}$ are the group means, $\overline{Y}_{i.}$ is the mean at the i^{th} level ($i = 1, 2$) of the covariate, $\overline{Y}_{.j}$ is the mean at the j^{th} level ($j = 1, 2$) of the treatment, and $\overline{Y}_{..}$ is the grand mean. The following possibilities arise.

1. The slopes of the lines joining the treatment means are parallel.



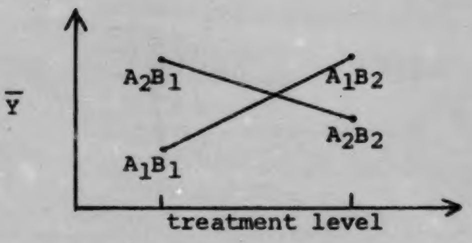
This is the case when the covariate, A, is working as desired; the difference between the treatment-level means is constant across the covariate levels. Either of the two quantities: $(\overline{Y}_{A_1B_2} - \overline{Y}_{A_1B_1})$ or $(\overline{Y}_{A_2B_2} - \overline{Y}_{A_2B_1})$ can be used to estimate the treatment effect.





2. The slopes of the lines joining the treatment means are moderately non-parallel. In this case, the "treatment effect" differs as a function of the covariate level, i.e, these two effects are correlated, positively or negatively, depending on the slope. However, in either case, the covariate lines do not meet within the feasible range of treatments. The covariate A can be used in the analysis, though its impact is to somewhat bias the treatment effects.

3. The slopes are excessively non-parallel. The lines of Y-means meet or cross within the feasibility range of treatment. This covariate cannot be used for any further analysis; however, this fact can be documented as an interesting research finding.



When the covariate and/or treatment is more than two-valued, similar analyses can be carried out, though with greater difficulty.

3.5 Study of Interactions

Fortunately, one does not need to discard an investigation just because a covariate happens to interact with some of the evaluation factors. In fact, even though the non-trivial interaction of a variable with an evaluation factor invalidates the former's claim to be a covariate in the study, the existence of an interaction can be a legitimate research finding. In section 3.4, we have illustrated how such an interaction can be interpreted and presented for a two-valued (potential) covariate. For the sake of generality, assume that the covariate is a quantitative variable, and that the research factor has only two levels, e.g., Follow Through and non-Follow Through. Let w denote the potential covariate, x (1 for Follow Through and 0 for non-Follow Through) denote the research factor and let $z = w \cdot x$ represent the interaction. Then the full model has the form:

$$\begin{aligned}\hat{Y} &= B_0 + B_w \cdot w + B_x \cdot x + B_z \cdot z \\ &= B_0 + B_w \cdot w + B_x \cdot x + B_z \cdot wx.\end{aligned}$$

This model uses three independent variables, and thus defines a three-dimensional hyperplane. If the signal associated with z is significant, i.e., if the F ratio:

$$F = \frac{n-4}{1} \cdot \frac{\hat{R}^2_{Y \cdot w,x,z} - \hat{R}^2_{Y \cdot w,x}}{1 - \hat{R}^2_{Y \cdot w,x,z}}$$

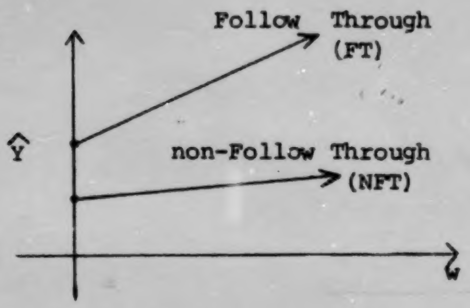
exceeds a predefined constant F_0, then w cannot serve as a covariate for this study. However, this research finding can be presented as follows. Note that the estimated values of the outcome measure, $\hat{Y}$, are given by:

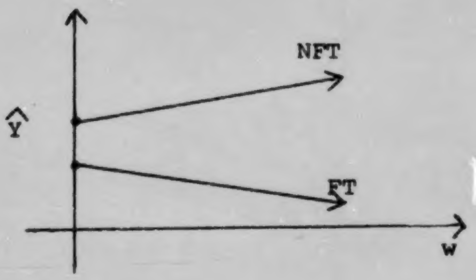
$$(21) \qquad \hat{Y} = (B_0 + B_x \cdot x) + (B_w + B_z \cdot x)w.$$

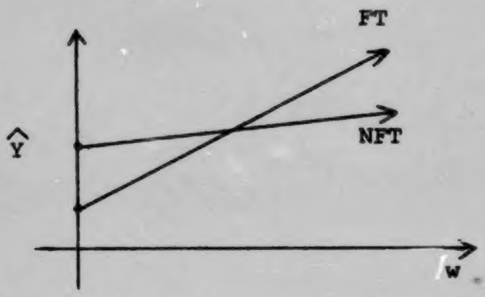
Let $\hat{Y}(1)$ denote the estimated value of Y when x = 1, i.e., for the group receiving Follow Through, and let $\hat{Y}(0)$ denote the corresponding values when x = 0, i.e. for the non-Follow Through group. From equation 21, one obtains:

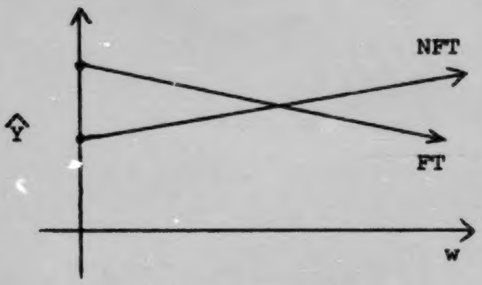
$$(22) \quad \begin{cases} \hat{Y}(0) = B_0 + B_w \cdot w \quad \text{and} \\ \hat{Y}(1) = (B_0 + B_x) + (B_w + B_z)w. \end{cases}$$

Notice that the interaction term, B_z, estimates the difference between the slopes of the predictor equations: this can be represented graphically as a pair of non-parallel straight lines, the actual slopes being determined by the data. As before, there are four possibilities: the Y-intercept of the Follow Through group can be above or below that of the non-Follow Through group, and B_z can be positive or negative. Though not a legitimate covariate, w provides valuable information about variations









in Y. This method of interpreting the interactions can be extended when there are more covariates and/or more research factors and/or more than two levels of each of the research factors. The method also applies when the covariate has non-linear as well as linear aspects.

4.0 ANALYTIC METHOD

The multiple regression format of ANOVA and ANCOVA was used to estimate adjusted and unadjusted effects within each of the studies. However, prior to the assessment of effects, it was necessary to select appropriate techniques for handling missing data, to develop coding schemes for the nonquantitative variables, and to explore the potentiality of heterogeneity of regression.

4.1 Missing Data

For those covariables which were deemed conceptually essential to the analysis and yet which were characterized by a small amount of missing data, the technique of employing a missing data code was adopted. Thus, each case contained both the value of the covariate and a missing data code. For a given case, if a valid response was present for the covariable, the missing data code variable was assigned a value of 0 to indicate presence of covariate information. If a valid response was not available, the appropriate design cell mean was entered as the covariate value, and the missing data code variable was assigned a value of 1 to indicate absence of covariate information. The proportion of variance accounted for by whether or not data are available could thus be ascertained. A test could then be made to determine if nonresponse was random with respect to the criterion variable. (See Cohen, 1973, for a detailed discussion of coding missing data.)

4.2 Coding of Nominal Variables

When a variable is qualitative or nominal in nature, i.e., when its values are used only for labels, it may be represented in the regression model by a set of g-1 predictor variables where g equals the number of mutually exclusive categories which characterize the nominal variable. Such variables may be nonsense, dummy, effects, or contrast coded. (See Volume I-B, chapter IV or Cohen, 1973, for a discussion of the particular benefits and costs of each technique.)

The variables which were nominally coded for the studies reported in this document are FT/NFT membership and "missingness" codes for Mother's

4

Education and family income. FT/NFT membership was effects coded +.5 for FT and -.5 for NFT. "Missingness" was dummy coded +1 to indicate that data for the variable was not available for the case and 0 to indicate presence of data for the variable for the case.

4.3 Exploration of Homogeneity (Heterogeneity of Regression)

The model and F test for exploring homogeneity/heterogeneity of regression are presented in Table MIII-1. The model included a single covariable (and its missingness code if it was a nominal variable), an effects coded variable representing FT/NFT membership, and the product of interaction of the covariable and FT/NFT membership. The model gives, in the interaction term, the within group, FT/NFT, deviations of the regression of the covariable on an outcome. A lack of deviation, nonsignificant F ratio, indicates homogeneous regression between FT and NFT as well as the appropriateness of the pooled regression. A significant F ratio indicates heterogeneous regression and further indicates that the within group coefficients deviate from each other substantially and cannot be pooled.

4.4 Determination of Unadjusted and Adjusted Effects

The ANOVA and ANCOVA models and their F tests for ascertaining the unadjusted and adjusted effects respectively are presented in Table MIII-1. The unadjusted effect is the raw score difference between the FT and NFT groups on Y, the outcome, that is, the difference between the group means. The adjusted effects represent the difference between groups taking into consideration the part of the outcome difference that correlates with differences on covariables. This adjusted difference represents the difference in gain. The adjusted difference is computed utilizing a model that includes the appropriate covariables and the FT/NFT membership variable. For the estimation of effect over the two year period from kindergarten to the end of first grade, the outcome needs to be adjusted for differences at entry to kindergarten. The covariable set for this two year effect thus includes Mother's Education plus missingness code and the fall kindergarten WRAT at the site level. Because of heterogeneity problems at the Sponsor level, the fall kindergarten WRAT was the only covariate used. For the estimation of effect in kindergarten, the outcome also needs to be adjusted

TABLE B-1.

F RATIOS AND FACTORS IN THE HIERARCHICAL MODEL

Factors:

Covariate = Fall kindergarten WRAT
Family Income
Mother's Education + Missingness Code

covset for effect = Mother's Education + Missingness Code (Missingness Code at site level only)
Fall kindergarten WRAT
Family Income (Sponsor level only)

covset for 2 year effect = Mother's Education + Missingness Code (site level only)
Fall kindergarten WRAT

A = FT/NFT

B = Covariate by FT/NFT

N = number of observations

k = number of factors = 1

Homogeneity/heterogeneity of regression:

$$sr_B^2 = R_Y^2.\text{cov A B} - R_Y^2.\text{cov B},$$

$$R_Y^2.\text{TOTAL} = R_Y^2.\ \text{cov A B}$$

$$F = \frac{sr_B^2}{(1 - R_Y^2.\text{TOTAL}) / (N - k - 1)}$$

ANOVA: Unadjusted Effect

$$sr_A^2 = R_Y^2.\text{A}$$

$$R_Y^2.\text{TOTAL} = R_Y^2$$

$$F = \frac{sr_A^2}{(1 - R_Y^2.\text{TOTAL}) / (N - k - 1)}$$

ANCOVA: Adjusted Effect

$$sr_B^2 = R_Y^2.\text{covset A} - R_Y^2\ \text{covset}$$

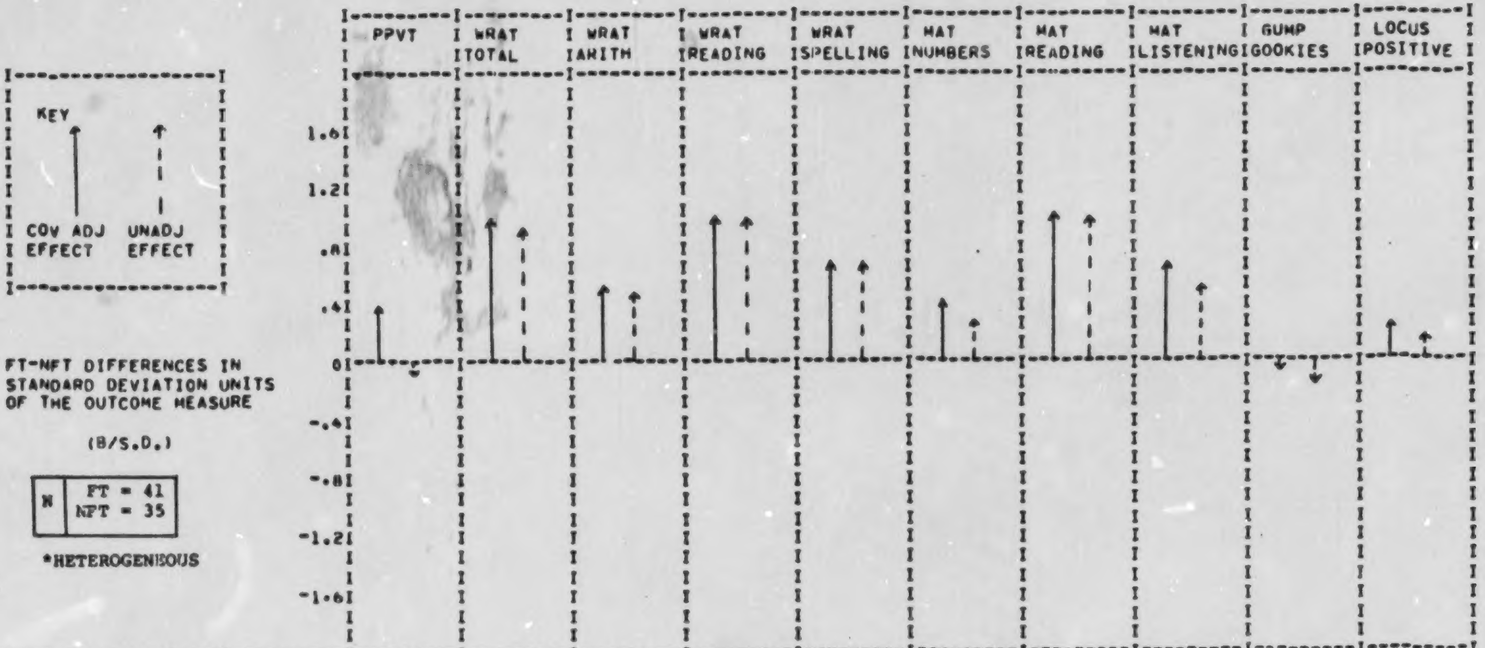
$$R_Y^2.\text{TOTAL} = R_Y^2.\text{covset A}$$

$$F = \frac{sr_B^2}{(1 - R_Y^2.\text{TOTAL}) / (N - k - 1)}$$

for differences at entry to kindergarten. The covariable set for the kindergarten effect thus included the fall kindergarten WRAT in addition to the Mother's Education plus missingness code at the site level. At the Sponsor level, Mother's Education, Family Income and the fall kindergarten WRAT were used as covariates. The details of these methods are presented in Table B-1.

APPENDIX C

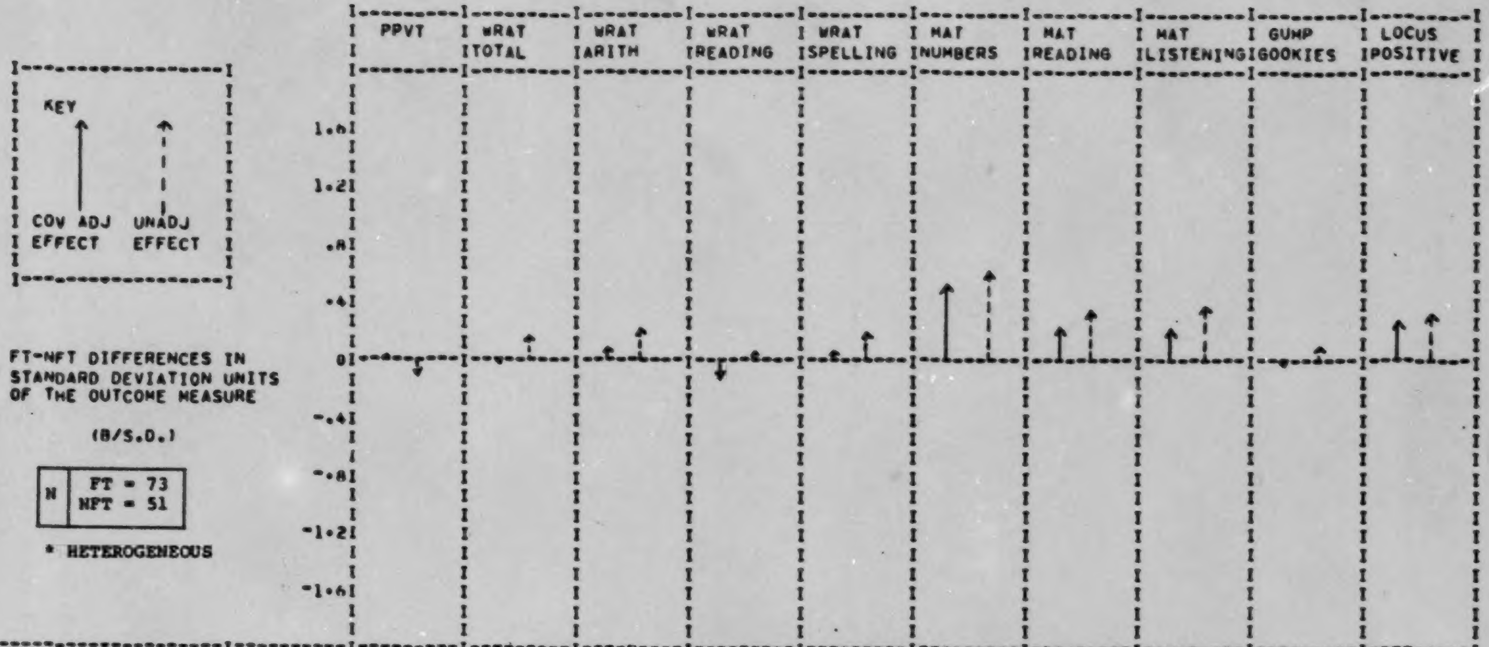
SITE LEVEL KINDERGARTEN EFFECTS PROFILE TABLES



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	2.635	8.956	1.573	* 5.576	1.807	2.232	5.298	4.046	-.009	.450
	UNADJ	-.668	8.440	1.446	5.192	1.80[illegible]	1.583	4.785	3.171	-.021	.263
S.E. = STANDARD ERROR OF B	ADJ	1.422	1.414	.653	.795	.605	1.138	.786	1.204	.031	.419
	UNADJ	1.611	1.860	.696	1.062	.613	1.235	.902	1.347	.030	.423
F= SIGNIFICANCE STATISTIC	ADJ	3.436	40.089	5.803	49.201	8.929	3.845	45.483	11.287	.076	1.154
	UNADJ	.172	20.584	4.314	23.904	8.627	1.643	28.128	5.539	.489	.388
SD=RAW SCORE STANDARD DEVIATION		6.962	9.077	3.091	5.273	2.797	5.391	4.575	6.028	.129	1.830
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.378	.987	.509	1.057	.646	.414	1.158	.671	-.066	.246
	UNADJ	-.096	.930	.468	.985	.644	.294	1.046	.526	-.162	.144

FIGURE C-1 KINDERGARTEN EFFECTS PROFILE FOR FAR WEST SITE NC2

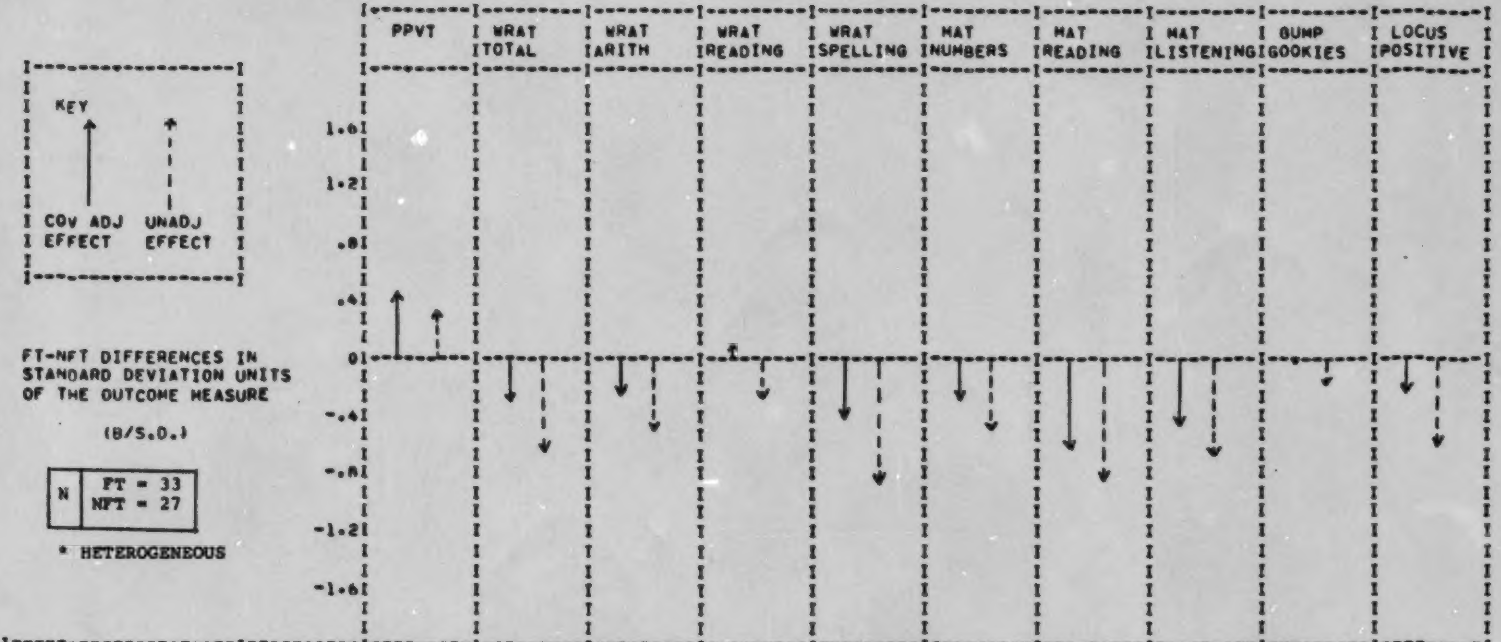
TABLE VIII-5



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	.062	-.128	.244	-.611	.239	3.549	1.026	1.465	-.002	.486
	UNADJ	-.636	1.673	.668	.152	.853	4.192	1.649	2.352	.013	.600
S.E. = STANDARD ERROR OF B	ADJ	.753	1.136	.448	.556	.643	.962	.716	.997	.003	.301
	UNADJ	1.135	1.869	.570	.800	.832	1.202	.863	1.209	.030	.335
F= SIGNIFICANCE STATISTIC	ADJ	.007	.013	.296	1.208	.138	13.618	2.055	2.159	.429	2.616
	UNADJ	.314	.802	1.373	.036	1.050	12.154	3.650	3.786	.186	3.215
SD=RAW SCORE STANDARD DEVIATION		6.153	10.144	3.101	4.332	4.525	6.829	4.741	6.644	.162	1.843
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.010	-.013	.079	-.141	.053	.520	.217	.221	-.012	.264
	UNADJ	-.103	.165	.215	.035	.189	.614	.348	.354	.079	.326

FIGURE C -2 KINDERGARTEN EFFECTS PROFILE FOR FAR WEST SITE NC1

FIGURE VIII-7



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

(B/S.D.)

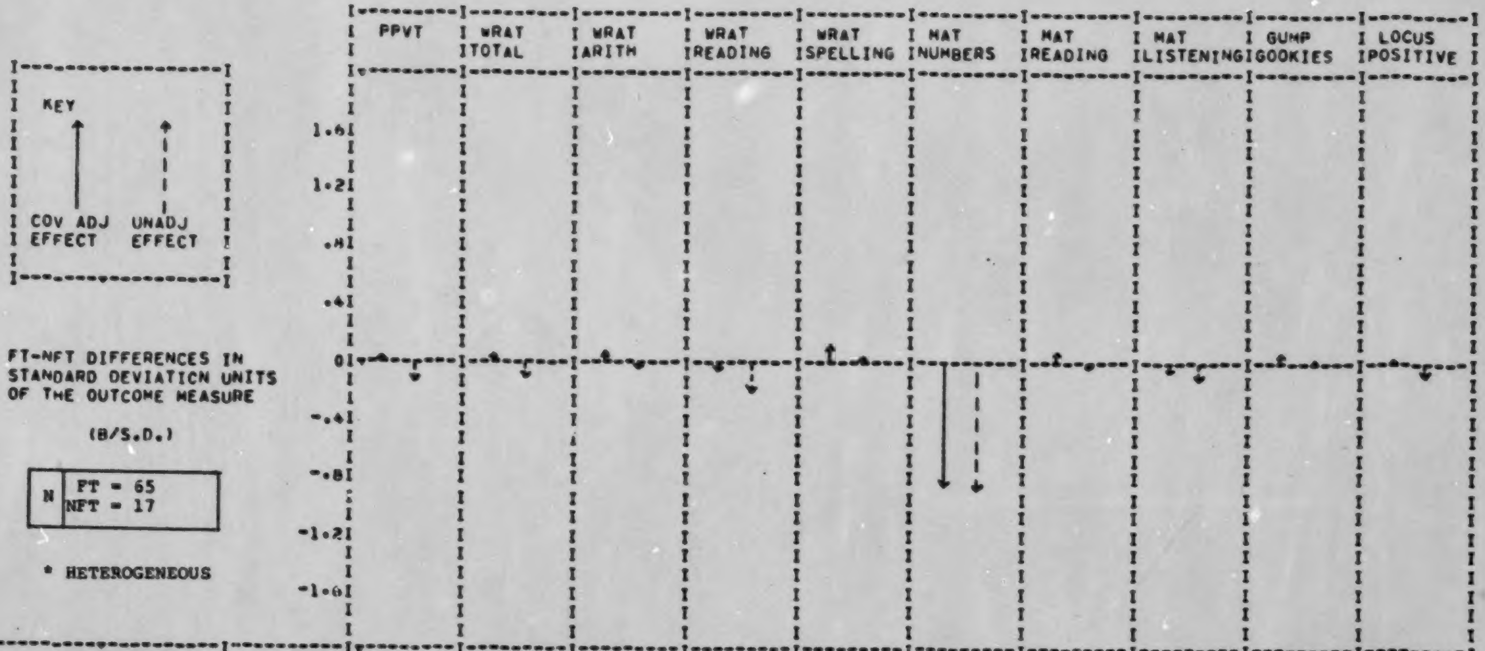
N	FT = 33 NFT = 27

* HETEROGENEOUS

		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	3.587	-3.612	-1.074	.410	-2.948	-1.815	-2.729	-2.394	-.001	-.691
	UNADJ	2.586	-7.970	-2.320	-1.485	-4.165	-2.946	-3.744	-3.519	-.024	-.997
S.E. = STANDARD ERROR OF B	ADJ	1.566	1.862	.880	.942	.885	1.296	.943	1.058	.037	.387
	UNADJ	2.005	3.010	1.144	1.313	1.109	1.510	1.047	1.277	.038	.410
F= SIGNIFICANCE STATISTIC	ADJ	5.247	3.763	1.487	.190	11.106	1.960	8.376	5.120	.000	3.196
	UNADJ	1.663	7.049	4.110	1.279	14.105	3.807	12.781	7.587	.399	5.900
SD=RAW SCORE STANDARD DEVIATION		7.770	12.177	4.524	5.072	4.724	5.956	4.421	5.190	.144	1.645
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.462	-.297	-.237	.081	-.624	-.305	-.617	-.461	-.005	-.420
	UNADJ	.333	-.654	-.513	-.293	-.882	-.495	-.847	-.678	-.165	-.606

FIGURE C-3 KINDERGARTEN EFFECTS PROFILE FOR FAR WEST SITE W1

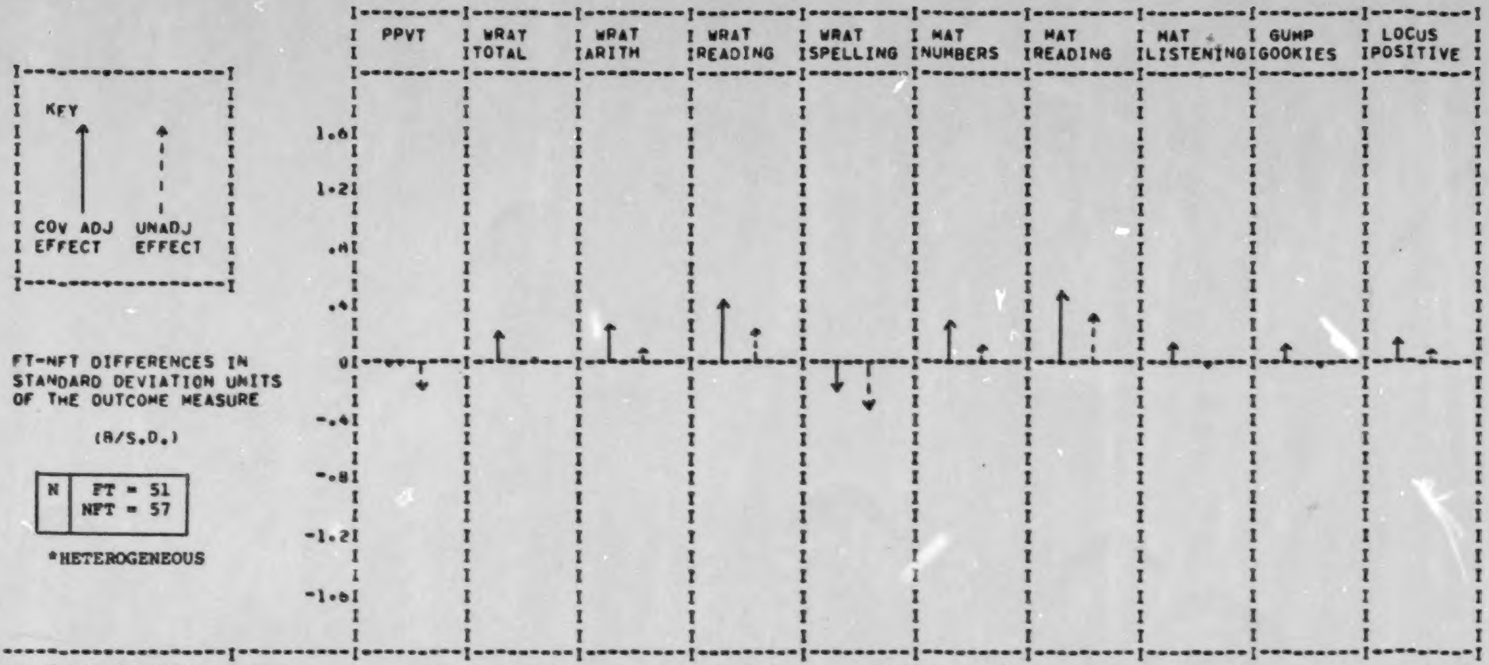
FIGURE VIII-9



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	.271	.546	.339	-.262	.470	-3.833	.263	-.287	.011	.034
	UNADJ	-1.373	-1.033	-.081	-1.071	.119	-3.905	-.077	-.531	.002	-.156
S.E. = STANDARD ERROR OF B	ADJ	2.701	2.132	.968	1.285	1.018	1.127	.977	1.601	.047	.607
	UNADJ	2.923	3.123	1.135	1.559	1.246	1.154	1.035	1.731	.046	.612
F= SIGNIFICANCE STATISTIC	ADJ	.010	.066	.122	.042	.213	11.576	.073	.032	.052	.003
	UNADJ	.221	.109	.005	.471	.009	11.453	.006	.094	.001	.065
SD=RAW SCORE STANDARD DEVIATION		10.680	11.402	4.142	5.706	4.546	4.501	3.776	6.318	.167	2.204
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.025	.048	.082	-.046	.103	-.852	.070	-.045	.065	.015
	UNADJ	-.129	-.091	-.020	-.188	.026	-.868	-.020	-.084	.010	-.071

FIGURE C-4 KINDERGARTEN EFFECTS PROFILE FOR FAR WEST SITE NE

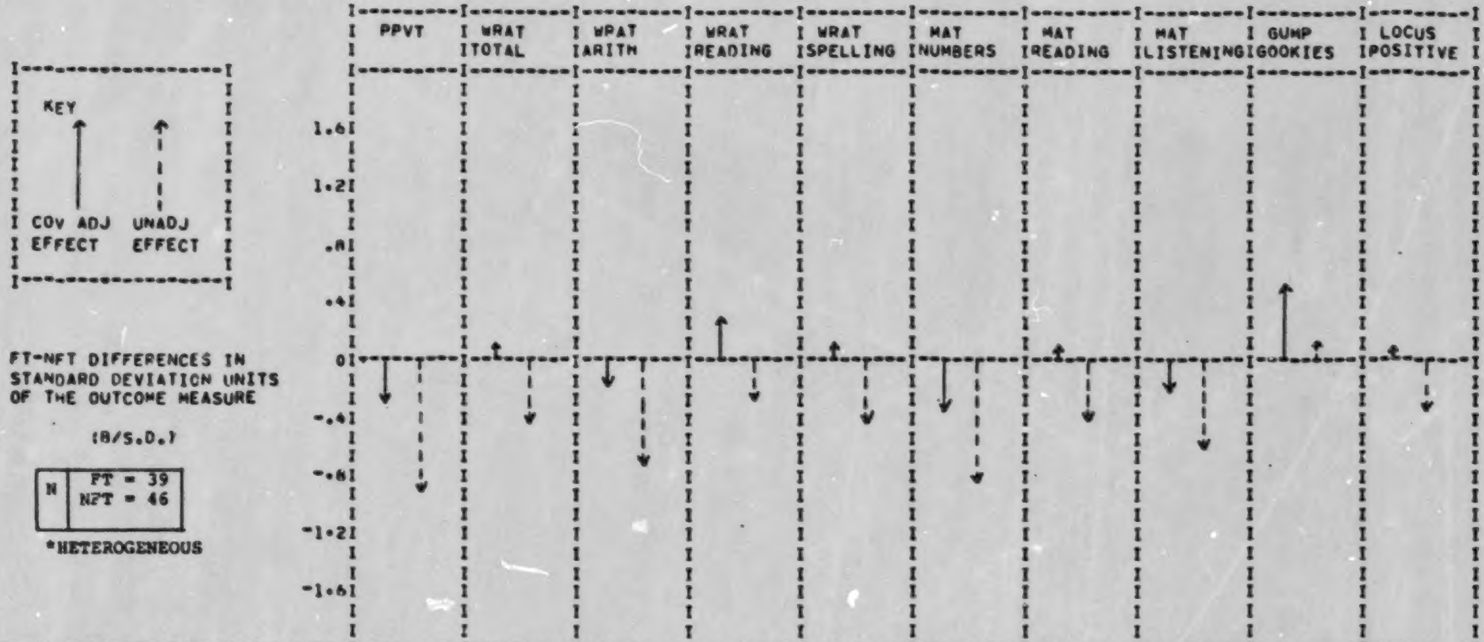
FIGURE VIII-1



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.024	1.972	.742	2.196	-.966	1.886	2.630	1.133	.011	.315
	UNADJ	-1.171	.050	.206	1.263	-1.420	.705	1.782	-.270	-.000	.140
S.E. = STANDARD ERROR OF B	ADJ	.841	1.328	.451	.784	.646	1.034	.839	1.378	.029	.364
	UNADJ	1.264	2.019	.572	1.026	.817	1.304	1.008	1.642	.030	.371
F= SIGNIFICANCE STATISTIC	ADJ	.001	2.203	2.709	7.844	2.235	3.328	9.828	.676	.133	.746
	UNADJ	.858	.001	.130	1.517	3.019	.292	3.128	.027	.000	.143
SD=RAW SCORE STANDARD DEVIATION		6.556	10.425	2.957	5.334	4.280	6.744	5.280	8.480	.153	1.916
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.004	.189	.251	.412	-.226	.280	.498	.134	.069	.164
	UNADJ	-.179	.005	.070	.237	-.332	.105	.338	-.032	-.001	.073

FIGURE C-5 KINDERGARTEN EFFECTS PROFILE FOR FAR WEST SITE W2

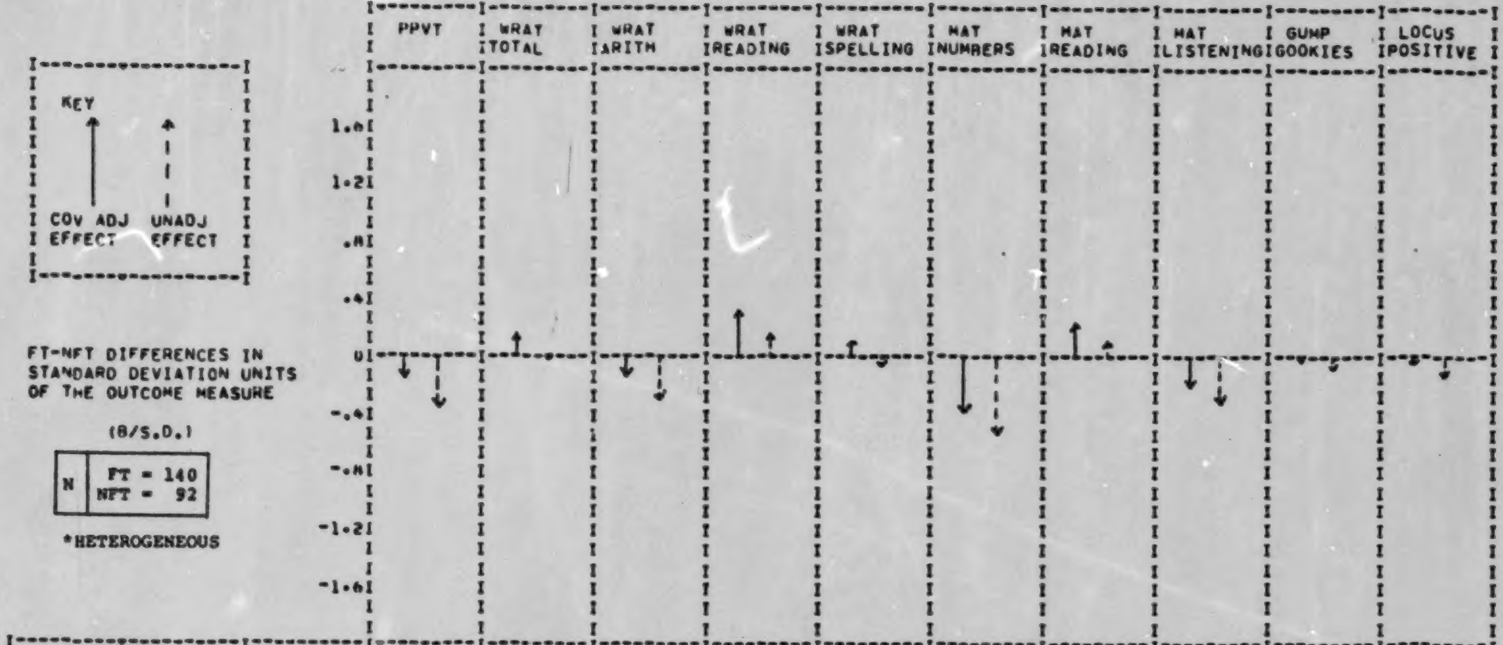
FIGURE VIII-15



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-2.347	1.207	-.643	1.406	* .445	-2.116	.493	-1.532	.074	.158
	UNADJ	-7.228	-6.230	-2.761	-1.438	-2.032	-5.214	-1.915	-4.521	.019	-.867
S.E. = STANDARD ERROR OF B	ADJ	1.331	1.648	.603	.961	.809	1.043	.823	1.418	.030	.495
	UNADJ	1.534	2.419	.767	1.138	.964	1.218	.953	1.490	.030	.512
F= SIGNIFICANCE STATISTIC	ADJ	3.107	.537	1.139	2.141	.302	4.114	.359	1.167	6.068	.101
	UNADJ	22.196	6.632	12.954	1.596	4.445	18.310	4.038	9.212	.407	2.868
SD=RAW SCORE STANDARD DEVIATION		7.851	11.416	3.747	5.216	4.491	6.118	4.432	7.131	.138	2.377
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.299	.106	-.172	.270	.099	-.346	.111	-.215	.536	.066
	UNADJ	-.921	-.546	-.737	-.276	-.452	-.852	-.432	-.634	.139	-.365

FIGURE C -6 KINDERGARTEN EFFECTS PROFILE FOR ARIZONA SITE S

FIGURE VIII-16



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

(B/S.D.)

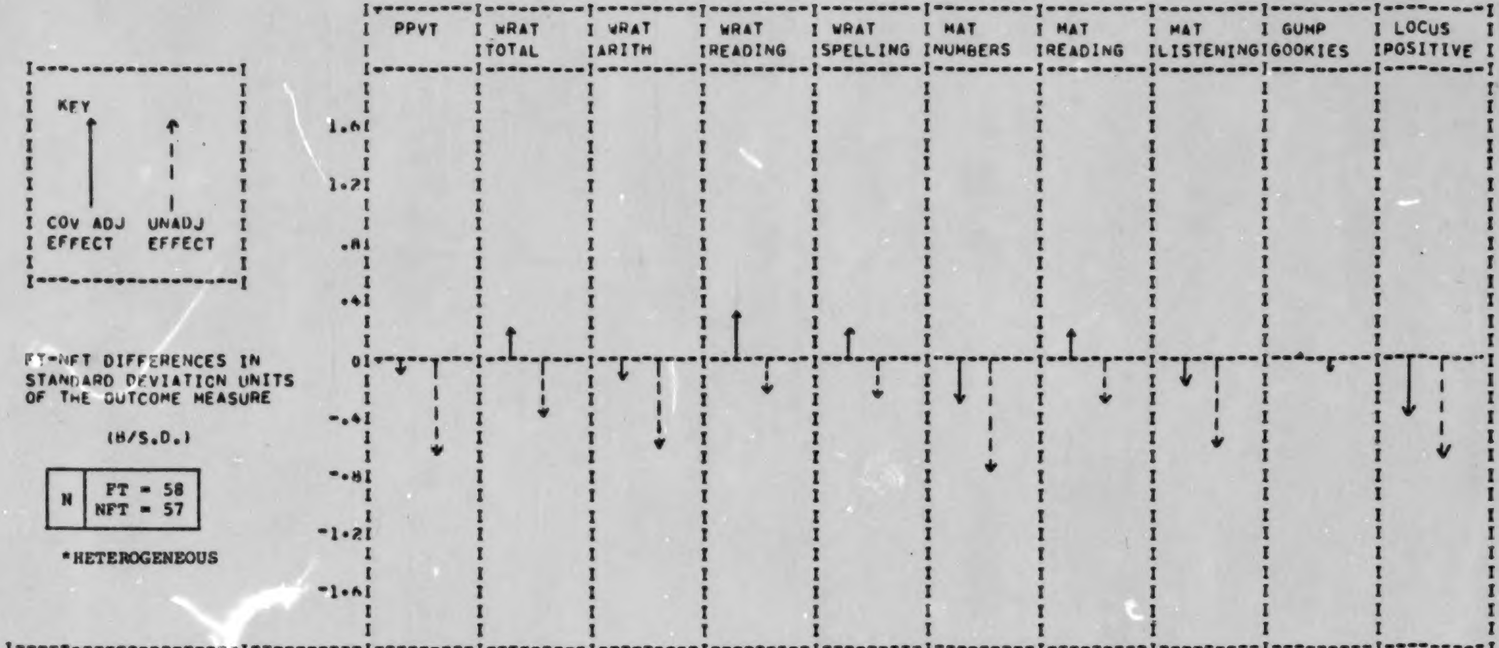
N	FT = 140 NFT = 92

*HETEROGENEOUS

		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.058	1.820	-.460	1.881	.399	-2.633	1.390	-1.515	-.003	-.068
	UNADJ	-2.429	-.365	-.958	.856	-.262	-3.610	.552	-2.600	-.010	-.205
S.E. = STANDARD ERROR OF B	ADJ	.610	.884	.319	.515	.466	.685	.552	.797	.021	.238
	UNADJ	.934	1.590	.436	.786	.627	.890	.770	1.066	.021	.250
F= SIGNIFICANCE STATISTIC	ADJ	3.003	4.240	2.084	13.322	.733	14.788	6.345	3.616	.015	.082
	UNADJ	6.764	.053	4.832	1.185	.174	16.464	.515	5.948	.212	.673
SD=RAW SCORE STANDARD DEVIATION		7.045	11.825	3.275	5.858	4.663	6.847	5.732	8.029	.158	1.856
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.150	.154	-.140	.321	.086	-.385	.243	-.189	-.016	-.037
	UNADJ	-.345	-.031	-.293	.146	-.056	-.527	.096	-.324	-.062	-.110

FIGURE C-7 KINDERGARTEN EFFECTS PROFILE FOR ARIZONA SITE NC1

FIGURE VIII-17



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

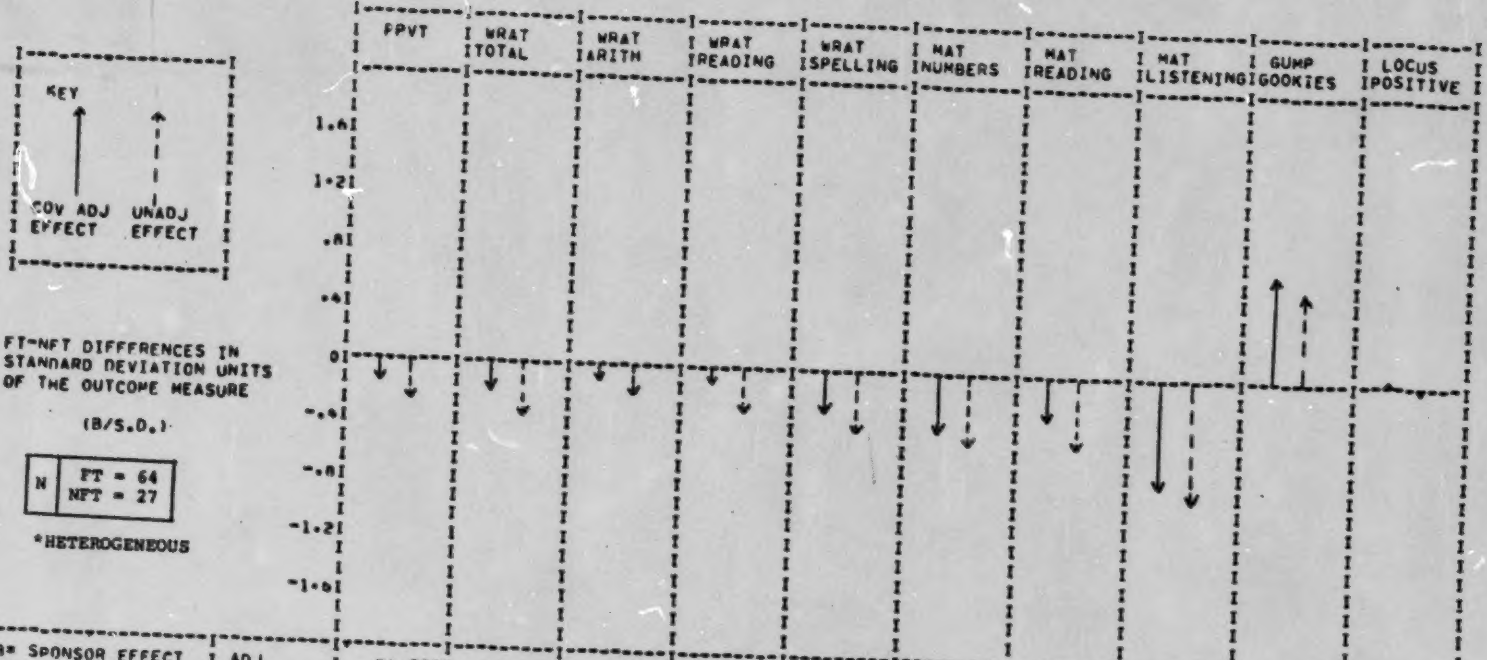
(B/S.D.)

N	FT = 58 NFT = 57

*HETEROGENEOUS

		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	* -.691	2.309	-.448	1.865	.893	* -1.970	1.071	* -1.319	.000	-.956
	UNADJ	-4.886	-4.426	-2.149	-1.246	-1.032	-5.115	-1.472	-4.393	-.010	-1.651
S.E. = STANDARD ERROR OF B	ADJ	.773	1.347	.520	.790	.643	.910	.814	1.135	.030	.426
	UNADJ	1.337	2.057	.633	1.042	.749	1.120	.948	1.314	.027	.418
F= SIGNIFICANCE STATISTIC	ADJ	.800	2.939	.743	5.575	1.925	4.684	1.734	1.351	.000	5.042
	UNADJ	13.349	4.629	11.527	1.430	1.895	20.851	2.412	11.175	.119	15.589
SD=RAW SCORE STANDARD DEVIATION		7.549	11.245	3.547	5.595	4.033	6.508	5.112	7.353	.147	2.382
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.092	.206	-.126	.333	.221	-.303	.210	-.179	.002	-.401
	UNADJ	-.647	-.395	-.606	-.223	-.256	-.786	-.288	-.597	-.065	-.693

FIGURE C-8 KINDERGARTEN EFFECTS PROFILE FOR ARIZONA SIZE MC2



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

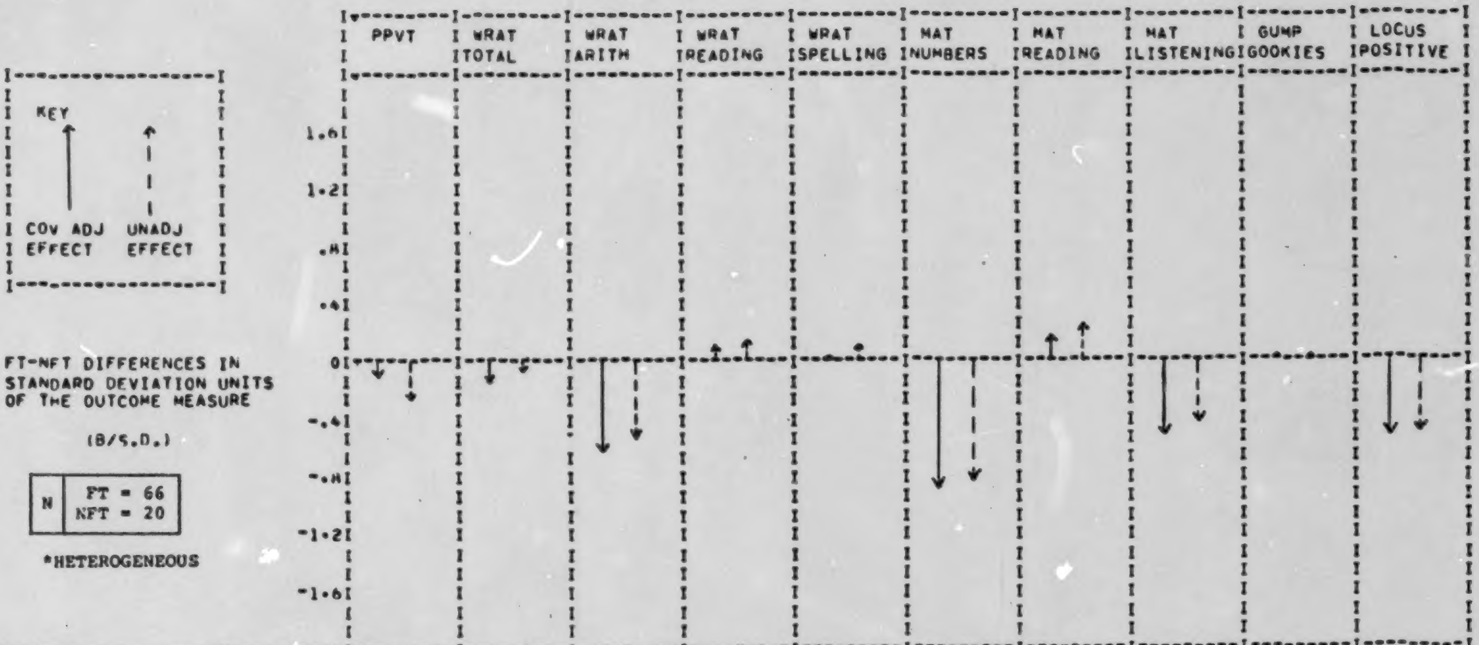
(B/S.D.)

N	FT = 64 NFT = 27

*HETEROGENEOUS

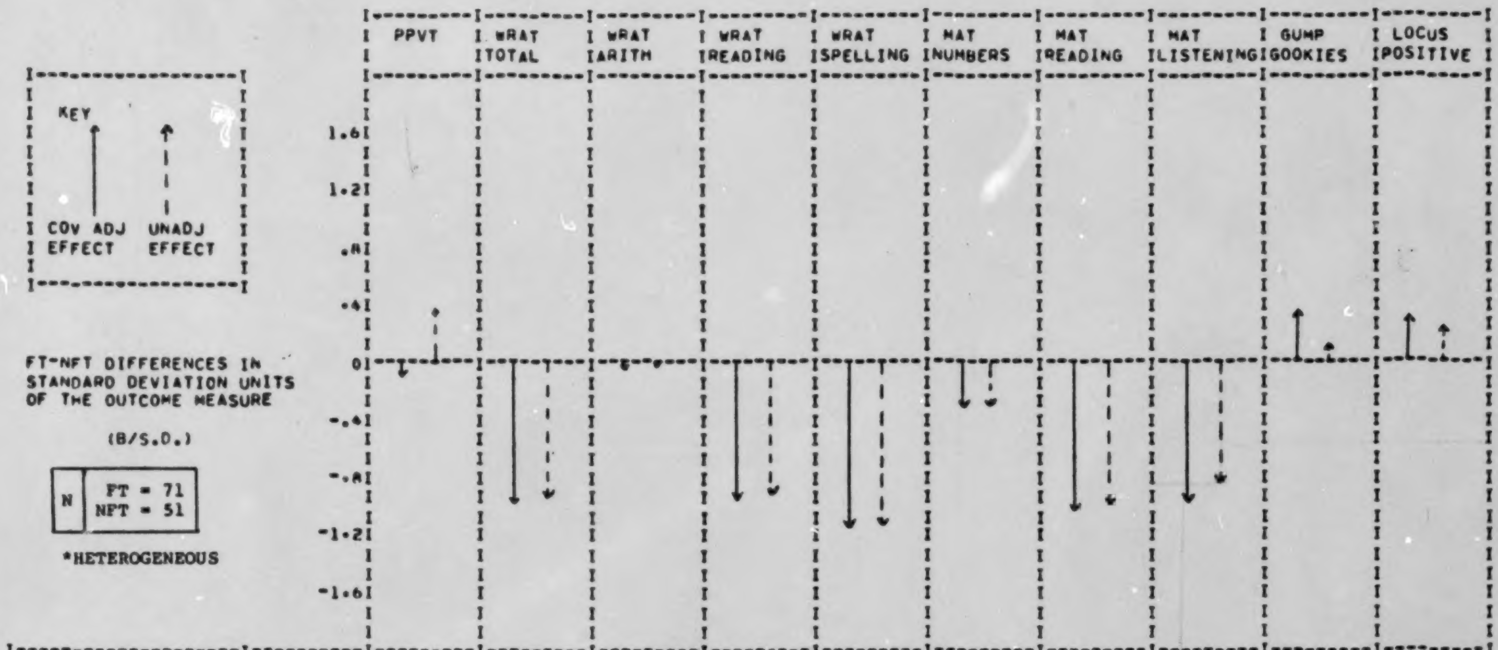
		FPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.529	-2.048	-.355	-.648	-1.045	-1.663	-1.447	-4.235	.107	.023
	UNADJ	-2.421	-3.766	-.631	-1.584	-1.551	-2.035	-2.288	-4.736	.093	-.072
S.E. = STANDARD ERROR OF B	ADJ	1.519	1.227	.491	.775	.726	.809	.928	1.065	.031	.523
	UNADJ	2.149	2.393	.737	1.215	.844	.947	1.093	1.223	.032	.516
F= SIGNIFICANCE STATISTIC	ADJ	1.013	2.784	.521	.700	2.070	4.226	2.433	15.811	11.634	.002
	UNADJ	1.269	2.674	.734	1.699	3.374	4.624	4.383	14.987	8.644	.019
SD=RAW SCORE STANDARD DEVIATION		9.380	10.130	3.206	5.315	3.728	4.207	4.852	5.731	.143	2.224
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.163	-.292	-.111	-.122	-.280	-.395	-.298	-.739	.747	.010
	UNADJ	-.258	-.372	-.197	-.298	-.416	-.484	-.472	-.826	.650	-.032

FIGURE C -9 KINDERGARTEN EFFECTS PROFILE FOR ARIZONA SITE NE1



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.147	-1.898	-2.451	.476	.077	-5.109	.909	-3.353	.002	-1.151
	UNADJ	-2.323	-.679	-2.127	1.070	.379	-4.782	1.344	-2.918	.004	-1.108
S.E. = STANDARD ERROR OF B	ADJ	2.081	2.076	.692	1.397	.939	1.199	.999	1.467	.038	.552
	UNADJ	2.432	3.322	.977	1.868	1.135	1.377	1.360	1.723	.038	.552
F= SIGNIFICANCE STATISTIC	ADJ	.304	.836	12.524	.116	.007	18.150	.829	5.228	.003	4.349
	UNADJ	.912	.042	4.736	.328	.111	12.050	.976	2.869	.013	4.029
SD=RAW SCORE STANDARD DEVIATION		9.524	12.941	3.913	7.288	4.425	5.737	5.327	6.824	.148	2.175
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.120	-.147	-.626	.065	.017	-.891	.171	-.491	.014	-.529
	UNADJ	-.244	-.052	-.544	.147	.086	-.834	.252	-.428	.030	-.509

FIGURE C.-10 KINDERGARTEN EFFECTS PROFILE FOR ARIZONA SITE NE2



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

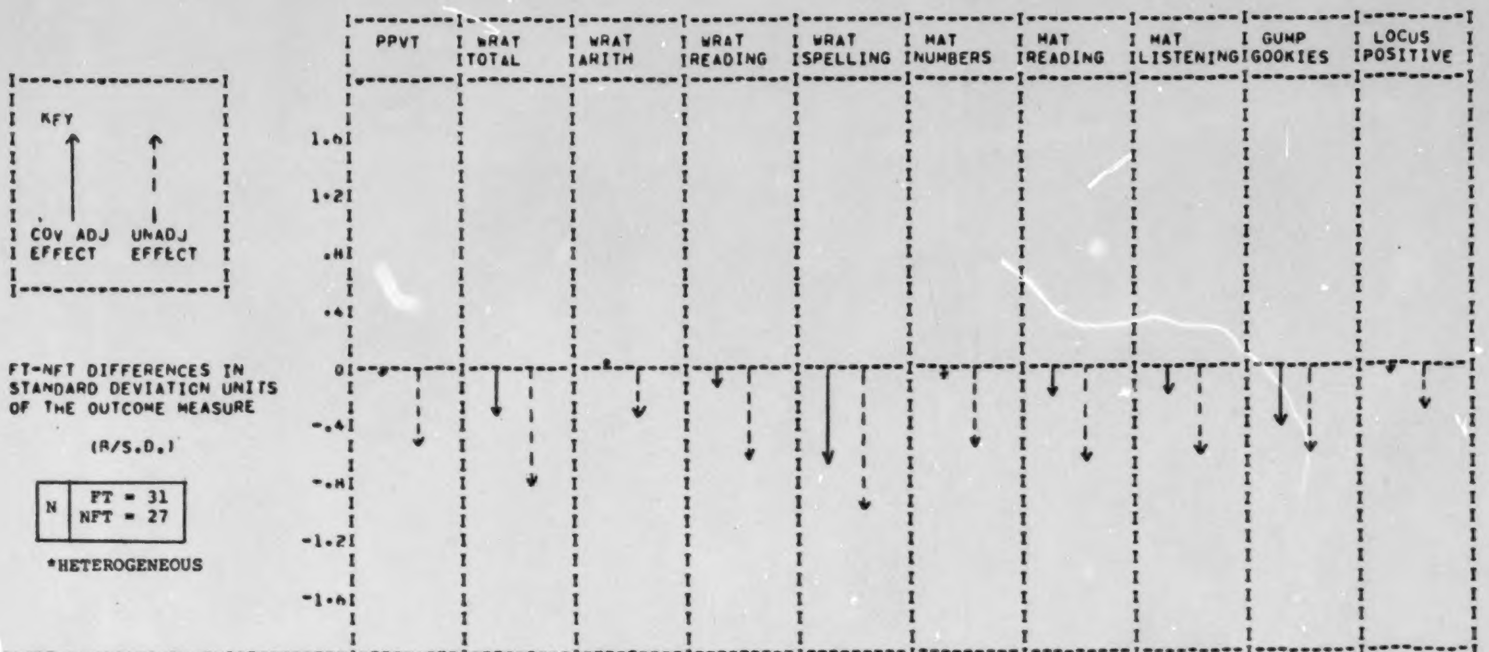
(B/S.D.)

N | FT = 71
NPT = 51

*HETEROGENEOUS

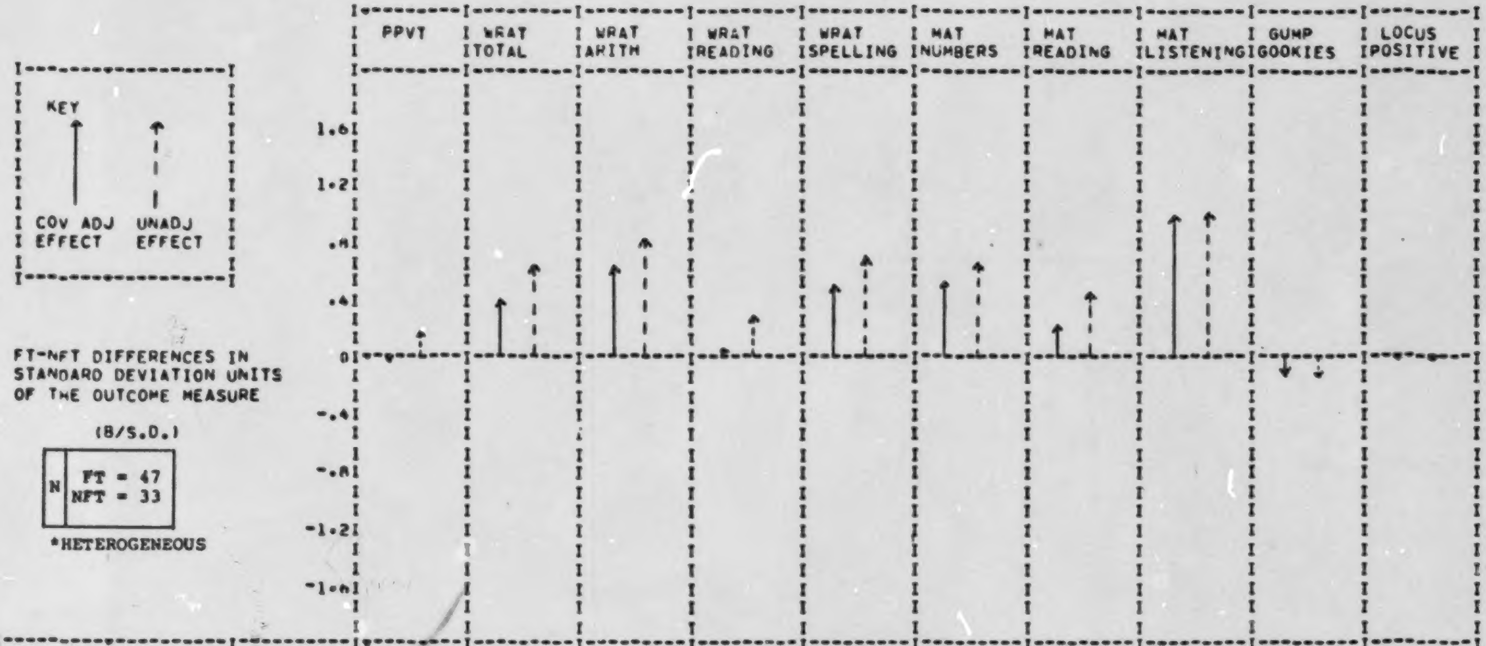
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.968	-10.931	-.145	-5.121	-5.665	-1.354	-5.583	-5.979	.058	.630
	UNADJ	3.789	-10.709	-.096	-5.007	-5.607	-1.257	-5.402	-5.280	.019	.498
S.E. = STANDARD ERROR OF B	ADJ	1.427	1.345	.504	.772	.659	.761	.839	1.020	.031	.387
	UNADJ	1.876	1.835	.609	.889	.751	.803	.875	1.072	.030	.373
F= SIGNIFICANCE STATISTIC	ADJ	.461	66.059	.083	44.056	73.929	3.164	44.291	34.396	3.424	2.653
	UNADJ	4.082	34.062	.025	31.692	55.751	2.449	38.121	24.262	.397	1.780
SD=RAW SCORE STANDARD DEVIATION		10.306	11.244	3.290	5.408	4.916	4.382	5.432	6.354	.163	2.022
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.094	-.972	-.044	- .947	-1.152	-.309	-1.028	-.941	.356	.312
	UNADJ	.368	-.952	-.029	- .926	-1.140	-.287	-.994	-.831	.117	.246

FIGURE C -11 KINDERGARTEN EFFECTS PROFILE FOR BANK STREET SITE NE2



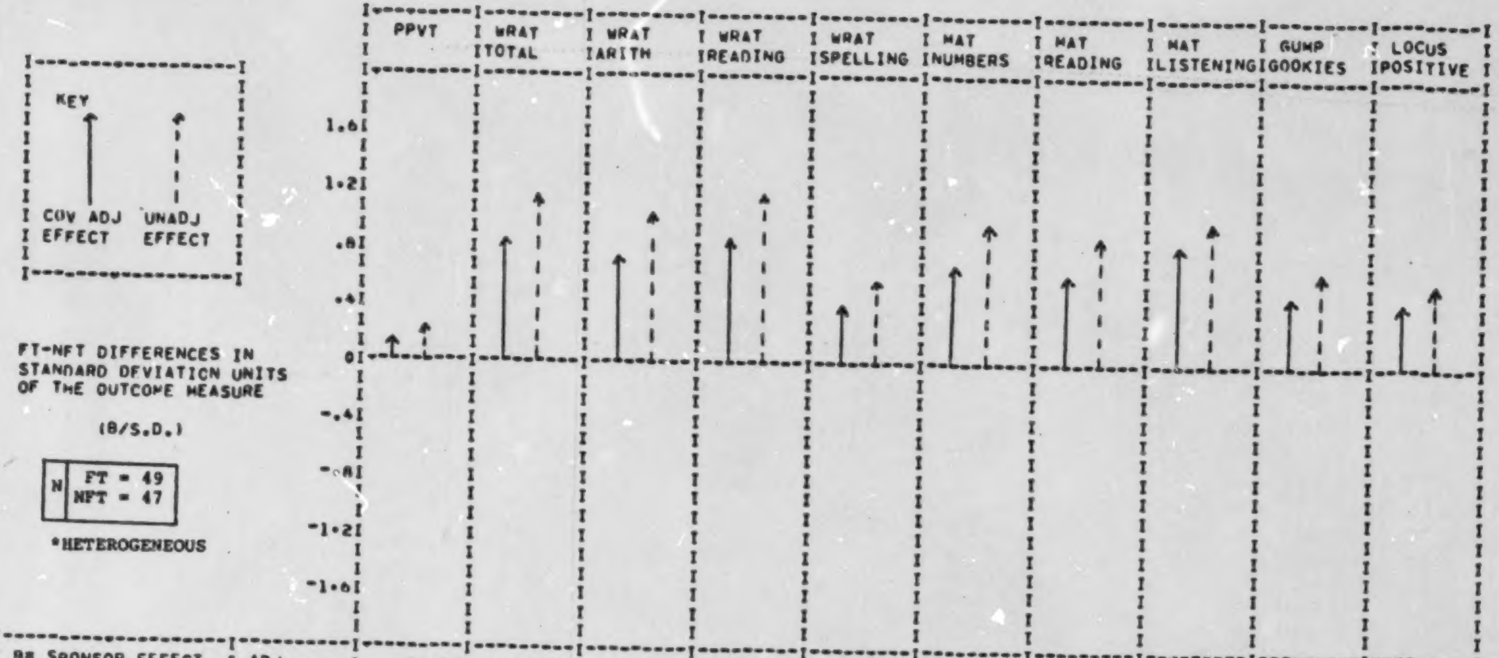
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.279	-3.523	.128	-.597	-3.055	-.278	-1.096	-1.310	-.071	-.112
	UNADJ	-3.278	-9.124	-.992	-3.565	-4.567	-2.683	-3.631	-4.928	-.104	-.559
S.E. = STANDARD ERROR OF B	ADJ	1.265	1.512	.631	.982	.870	.905	1.125	1.528	.048	.543
	UNADJ	1.617	2.720	.787	1.411	1.085	1.237	1.403	1.996	.045	.553
F= SIGNIFICANCE STATISTIC	ADJ	.047	5.424	.041	.370	12.328	.094	.950	.735	2.203	.042
	UNADJ	4.113	11.256	1.586	6.387	17.728	4.704	6.698	6.095	5.437	1.021
SD=RAW SCORE STANDARD DEVIATION		6.307	11.222	3.006	5.606	4.687	4.850	5.589	7.915	.176	2.102
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.044	-.314	.043	-.107	-.652	-.057	-.196	-.166	-.406	-.053
	UNADJ	-.520	-.813	-.330	-.636	-.975	-.553	-.650	-.623	-.591	-.266

FIGURE C-12 KINDERGARTEN EFFECTS PROFILE FOR BANK STREET SITE NE1



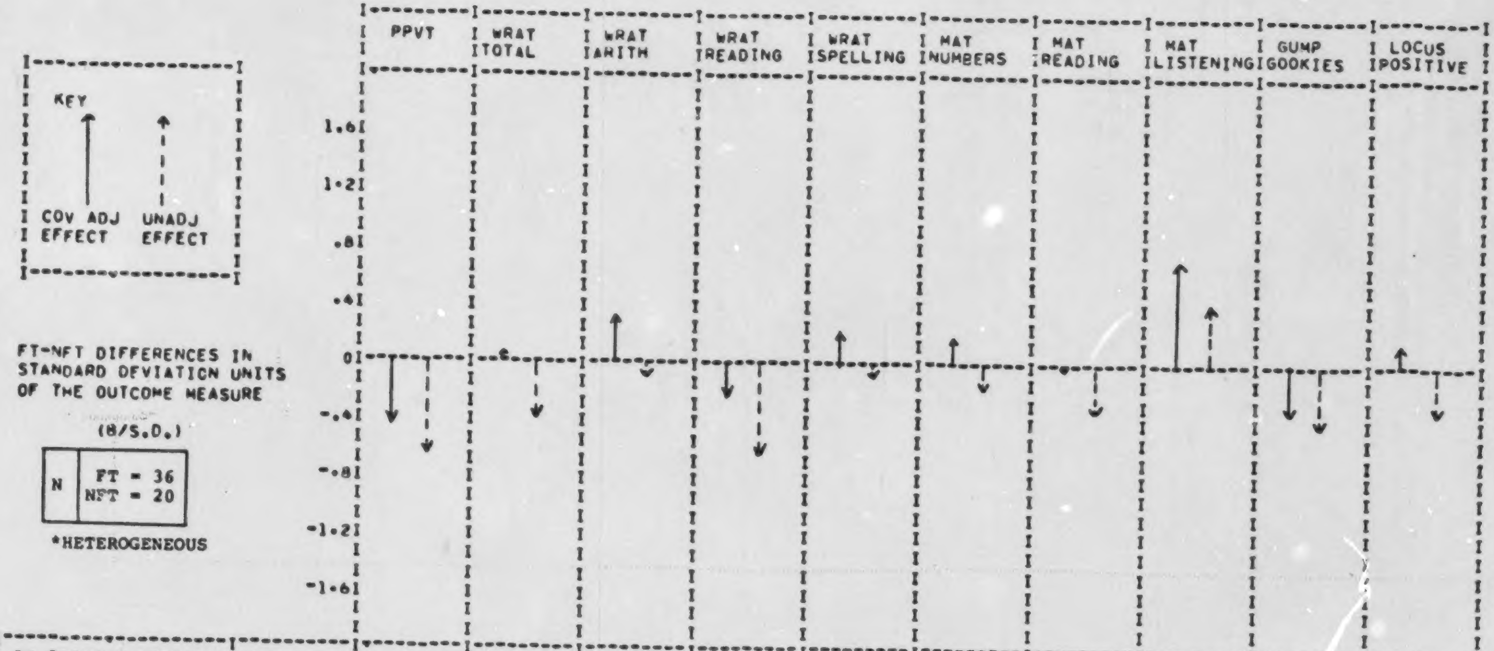
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.214	4.645	2.196	.339	2.110	3.513	1.216	7.812	-.046	.144
	UNADJ	1.183	7.788	2.778	1.981	3.029	4.364	2.451	8.499	-.046	.333
S.E. = ST[illegible] ERROR OF B	[illegible]	1.263	1.949	.620	1.338	.785	1.269	1.013	1.461	.037	.447
	UNADJ	1.62[illegible]	2.079	.712	1.608	.927	1.419	1.190	1.520	.038	.441
F= SIGNIFICANCE STATISTIC	ADJ	.029	5.679	12.552	.064	7.224	7.658	1.442	28.604	1.546	.103
	UNADJ	.528	8.451	15.227	1.518	10.674	9.455	4.240	31.273	1.500	.569
SD=RAW SCORE STANDARD DEVIATION		7.147	12.340	3.405	7.104	4.325	6.575	5.347	7.870	.167	1.936
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.030	.376	.645	.048	.488	.534	.227	.993	-.275	.074
	UNADJ	.166	.631	.816	.279	.700	.664	.458	1.080	-.277	.172

FIGURE C-13 KINDERGARTEN EFFECTS PROFILE FOR OREGON SITE NC1



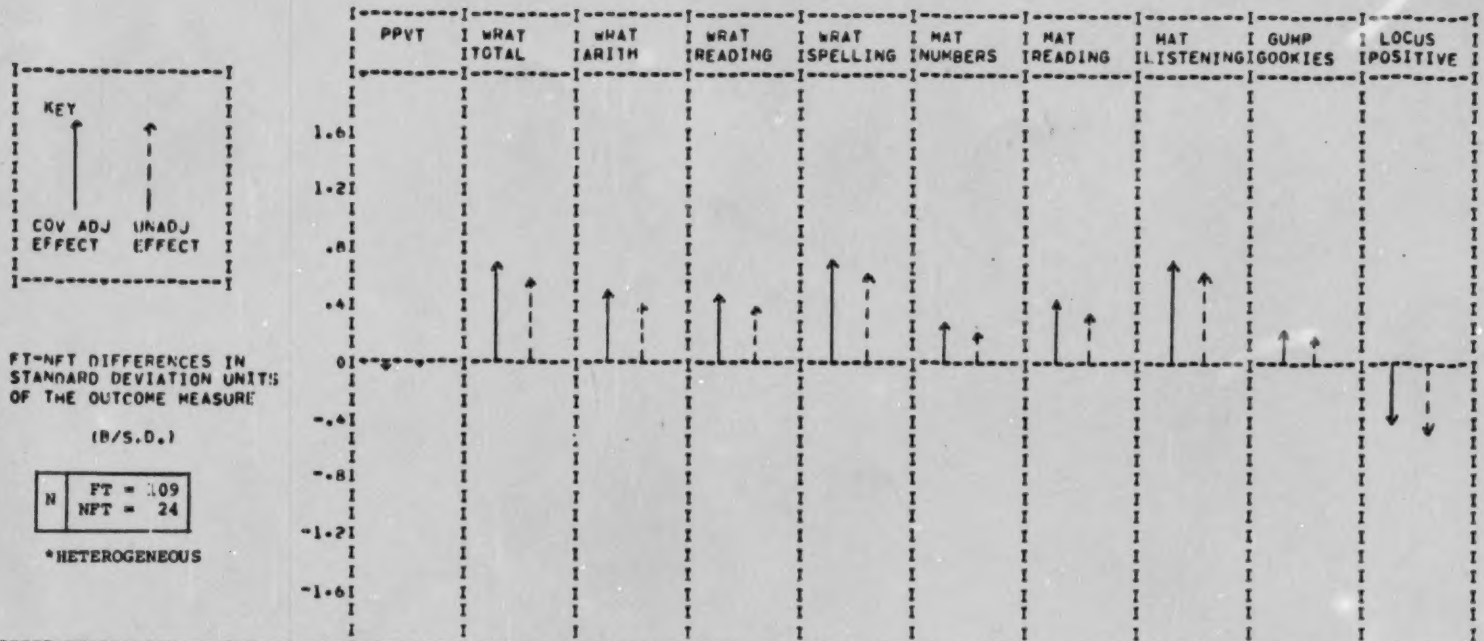
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	1.446	11.425	2.535	6.817	2.073	4.814	3.917	5.634	.071	.828
	UNADJ	2.337	15.744	3.619	9.038	3.087	7.059	5.557	6.964	.099	1.139
S.E. = STANDARD ERROR OF B	ADJ	1.801	1.877	.513	1.085	1.002	1.091	.944	1.161	.029	.374
	UNADJ	1.969	2.256	.598	1.284	.999	1.242	1.134	1.185	.029	.362
F= SIGNIFICANCE STATISTIC	ADJ	.645	37.052	24.422	39.489	4.281	19.467	17.227	23.534	5.936	4.908
	UNADJ	1.407	48.688	36.621	49.580	9.546	29.866	24.031	34.543	11.644	9.885
SD=RAW SCORE STANDARD DEVIATION		9.667	13.544	3.434	7.729	5.109	7.224	6.189	6.751	.151	1.856
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.150	.844	.738	.882	.406	.666	.633	.835	.471	.446
	UNADJ	.242	1.162	1.054	1.169	.604	.977	.898	1.032	.660	.614

FIGURE . C-14 KINDERGARTEN EFFECTS PROFILE FOR OREGON SITE NC2



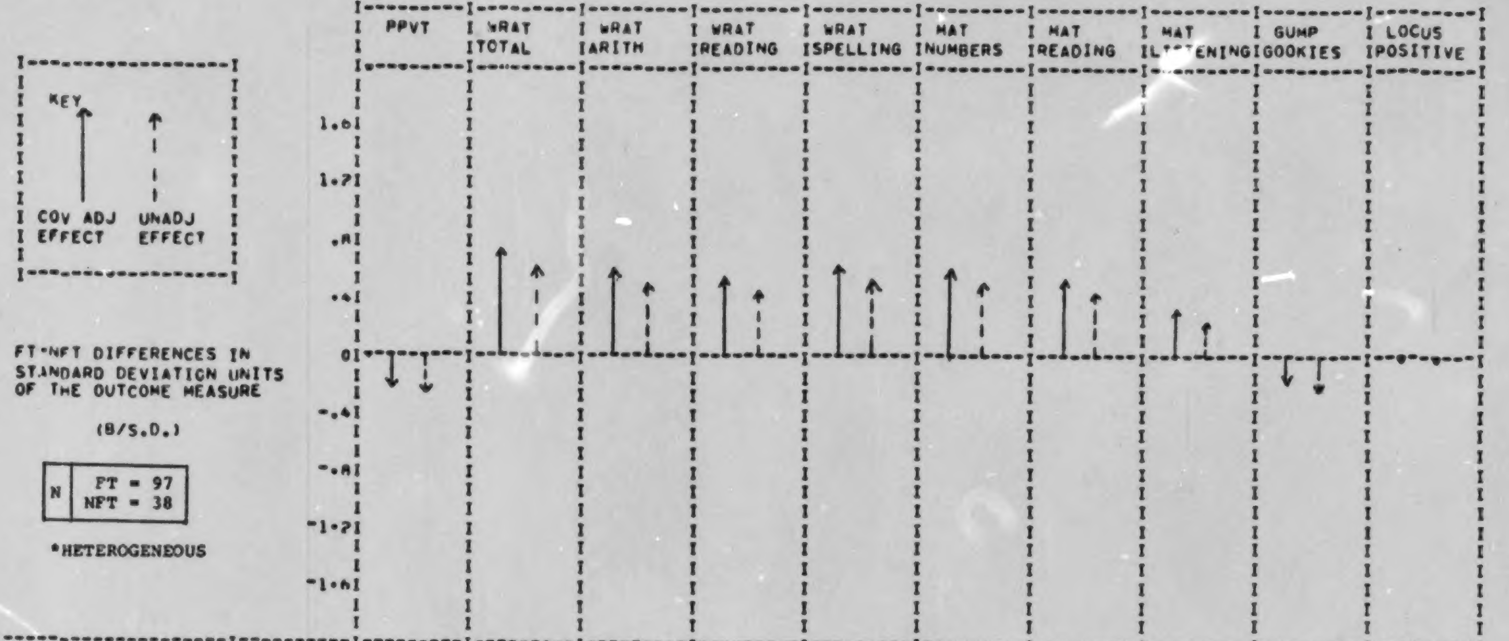
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-3.788	.661	1.140	-1.454	.975	1.284	-.065	4.828	-.047	.316
	UNADJ	-5.072	-4.761	-.356	-4.050	-.356	-1.250	-1.983	2.856	-.057	-.667
S.E. = STANDARD ERROR OF B	ADJ	2.025	2.543	.902	1.535	1.195	1.935	1.572	1.786	.041	.578
	UNADJ	2.216	3.292	1.014	1.719	1.260	1.969	1.680	1.821	.038	.603
F= SIGNIFICANCE STATISTIC	ADJ	3.499	.068	1.596	.897	.666	.440	.002	7.305	1.294	.299
	UNADJ	5.241	2.092	.123	5.550	.080	.403	1.393	2.458	2.238	1.222
SD=RAW SCORE STANDARD DEVIATION		8.245	11.921	3.607	6.414	4.481	7.022	6.047	6.617	.138	2.167
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.459	.055	.316	-.227	.218	.183	-.011	.730	-.341	.146
	UNADJ	-.615	-.399	-.099	-.631	-.079	-.178	-.328	.432	-.412	-.308

FIGURE C-15 KINDERGARTEN EFFECTS PROFILE FOR OREGON SITE NE



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.594	8.092	2.015	3.004	3.073	1.993	2.093	5.009	.036	-.805
	UNADJ	-.358	6.652	1.626	2.342	2.684	1.376	1.705	4.468	.027	-.902
S.E. = STANDARD ERROR OF B	ADJ	1.639	1.922	.744	1.141	.822	1.525	.968	1.322	.035	.423
	UNADJ	2.131	2.587	.892	1.381	.933	1.679	1.070	1.515	.036	.432
F= SIGNIFICANCE STATISTIC	ADJ	.131	17.737	7.343	6.931	13.987	1.707	4.675	14.361	1.052	3.624
	UNADJ	.028	4.611	3.321	2.876	8.270	.672	2.538	8.698	.597	4.365
SD=RAW SCORE STANDARD DEVIATION		9.378	11.673	3.977	6.144	4.237	7.408	4.757	6.888	.154	1.903
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.063	.693	.507	.489	.725	.269	.440	.727	.232	-.423
	UNADJ	-.036	.570	.409	.381	.634	.186	.358	.649	.178	-.474

FIGURE C -16 KINDERGARTEN EFFECTS PROFILE FOR KANSAS SITE NE



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

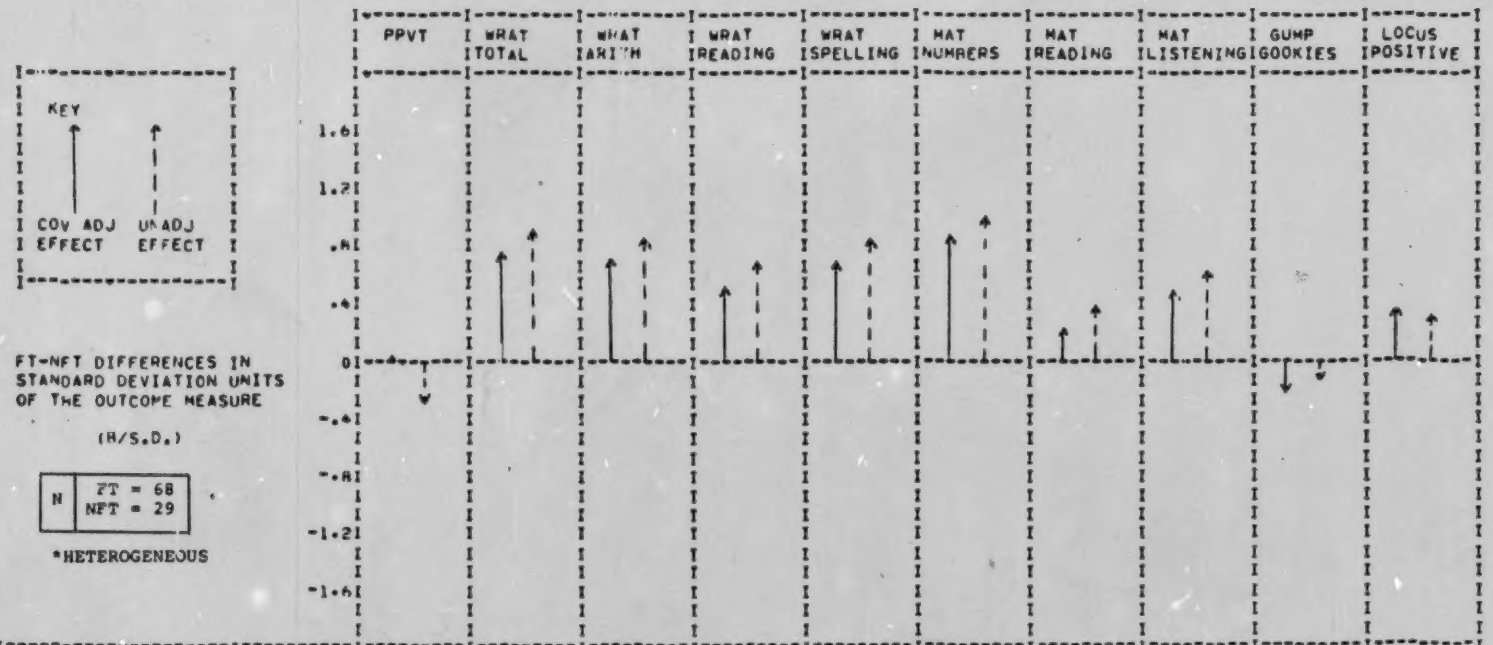
(B/S.D.)

N	FT = 97 NFT = 38

*HETEROGENEOUS

		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.828	7.590	2.234	2.678	2.678	3.106	2.646	2.533	-.031	-.078
	UNADJ	-2.023	6.326	1.782	2.239	2.307	2.689	2.234	1.848	-.041	-.078
S.E. = STANDARD ERROR OF B	ADJ	1.120	1.604	.593	.849	.746	.892	.874	1.449	.033	.438
	UNADJ	1.468	1.967	.684	.942	.798	.980	.935	1.519	.032	.419
F= SIGNIFICANCE STATISTIC	ADJ	2.666	22.392	14.209	9.954	12.885	12.111	9.157	3.055	.869	.032
	UNADJ	1.899	10.353	6.[illegible]77	5.657	8.361	7.532	5.706	1.480	1.636	.035
SD=RAW SCORE STANDARD DEVIATION		7.697	10.630	3.653	5.004	4.283	5.243	4.971	7.951	.163	2.141
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.238	.714	.611	.535	.625	.592	.532	.319	-.192	-.036
	UNADJ	-.263	.595	.488	.447	.539	.513	.449	.232	-.249	-.036

FIGURE C-17 KINDERGARTEN PROFILE FOR KANSAS SITE MC2

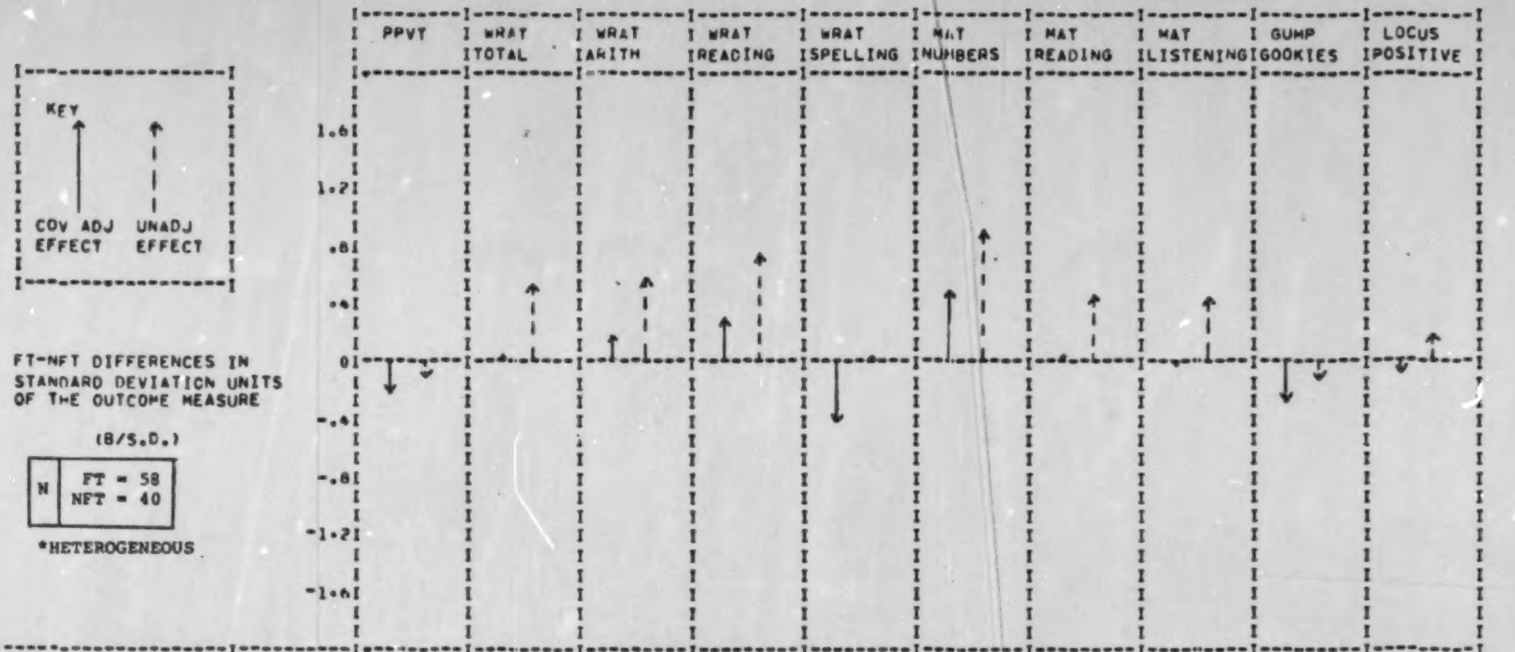


N	
	FT = 68 NFT = 29

*HETEROGENEOUS

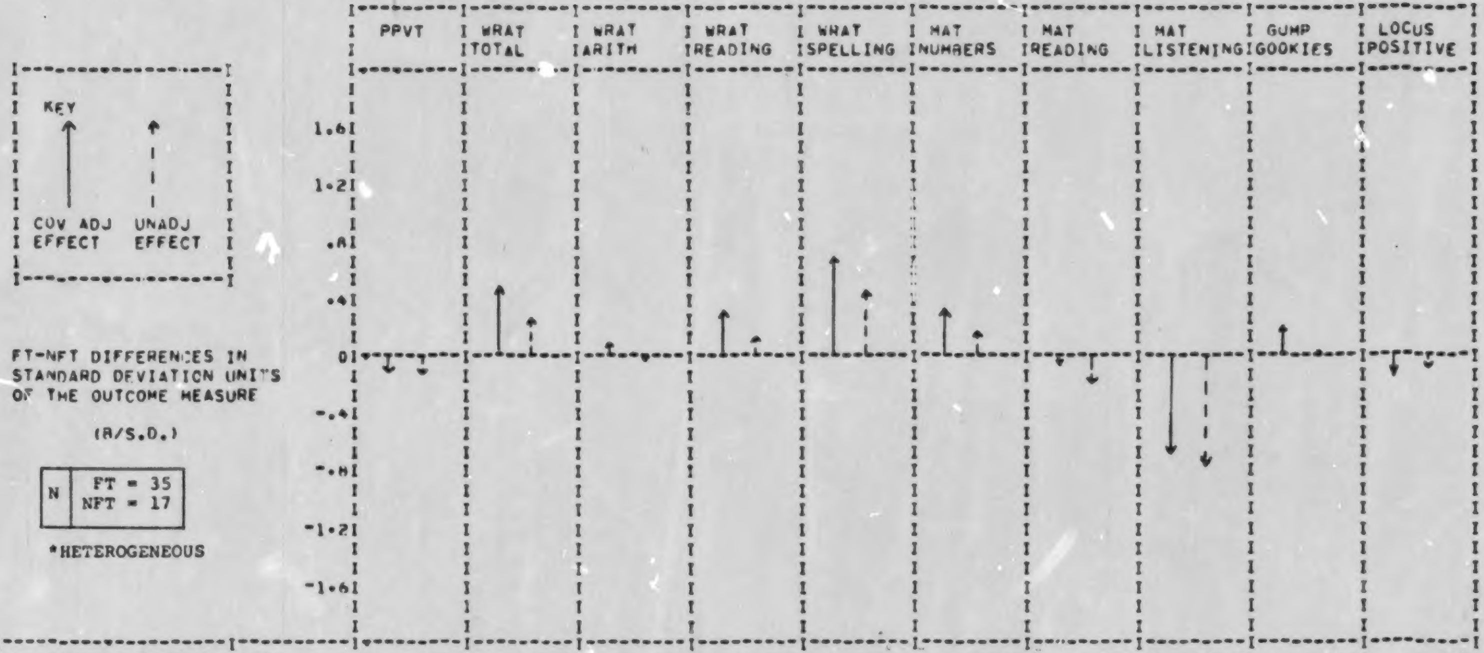
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	.345	10.109	3.350	3.101	3.658	5.785	1.188	* 3.614	* -.034	.745
	UNADJ	-2.345	12.464	4.055	3.945	4.464	6.725	1.905	4.750	-.020	.718
S.E. = STANDARD ERROR OF B	ADJ	1.347	1.980	.798	.966	.931	.973	.995	1.297	.033	.494
	UNADJ	1.800	2.798	.990	1.261	1.083	1.313	1.177	1.618	.033	.486
F= SIGNIFICANCE STATISTIC	ADJ	.065	26.053	17.644	10.294	15.448	35.351	1.428	7.759	1.081	2.273
	UNADJ	1.697	19.839	16.776	9.790	17.009	26.247	2.618	8.616	.358	2.180
SD=RAW SCORE STANDARD DEVIATION		8.102	13.740	4.796	5.911	5.249	6.624	5.326	7.543	.149	2.204
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.043	.730	.699	.525	.697	.873	.223	.479	-.229	.338
	UNADJ	-.289	.907	.846	.667	.851	1.015	.358	.630	-.132	.325

FIGURE C-18 KINDERGARTEN EFFECTS PROFILE FOR KANSAS SITE S



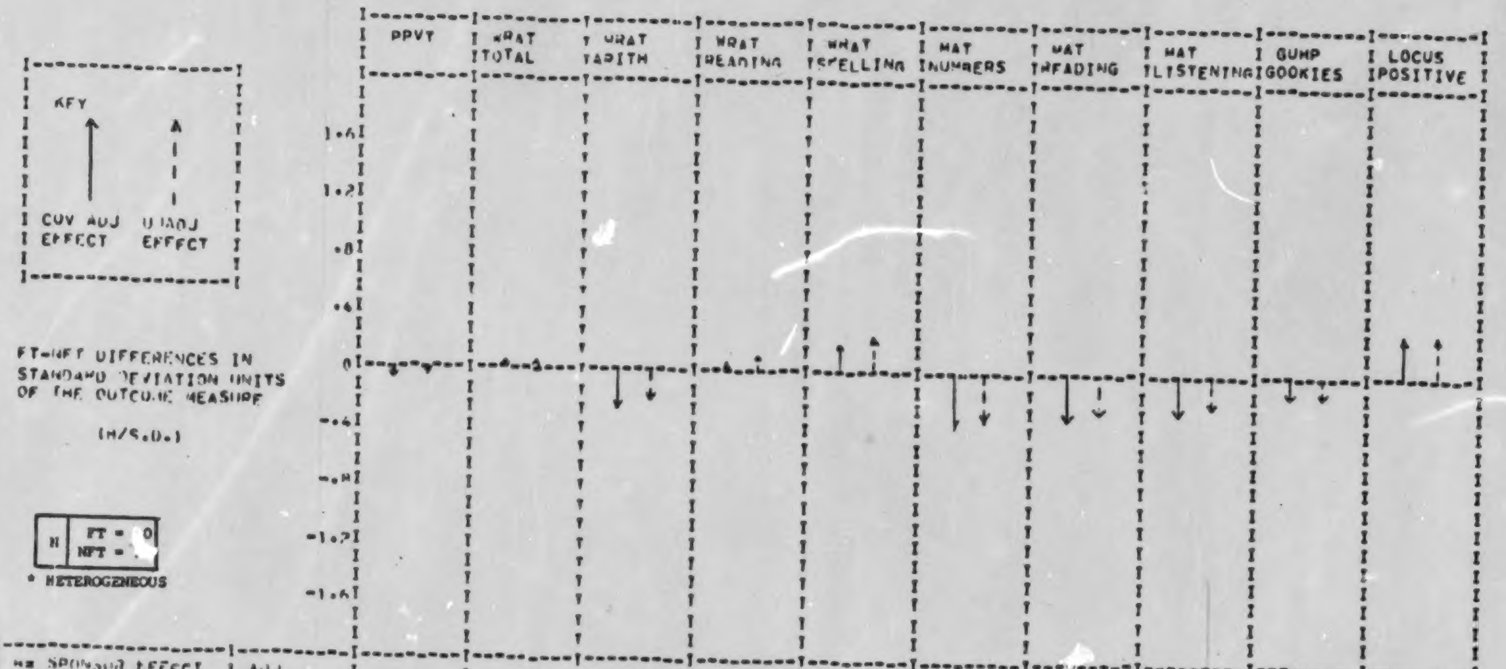
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.775	.502	.672	1.654	-1.825	3.383	.206	-.056	-.042	-.201
	UNADJ	-.805	4.647	2.484	4.153	.010	6.361	2.372	2.843	-.019	.367
S.E. = STANDARD ERROR OF B	ADJ	.977	1.846	.701	.863	.803	1.023	.909	1.075	.031	.407
	UNADJ	1.745	1.505	.850	1.113	.924	1.305	1.057	1.324	.030	.433
F= SIGNIFICANCE STATISTIC	ADJ	3.297	.077	.919	3.674	5.167	10.923	.051	.003	1.804	.243
	UNADJ	.213	7.040	8.544	13.923	.000	23.761	5.038	4.613	.407	.719
SD=RAW SCORE STANDARD DEVIATION		8.454	12.562	4.293	5.764	4.473	7.055	5.249	6.559	.146	2.104
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.210	.040	.157	.287	-.408	.479	.039	-.009	-.287	-.095
	UNADJ	-.095	.529	.579	.720	.002	.902	.452	.433	-.132	.175

FIGURE C-19 KINDERGARTEN EFFECTS PROFILE FOR KANSAS SITE NC1



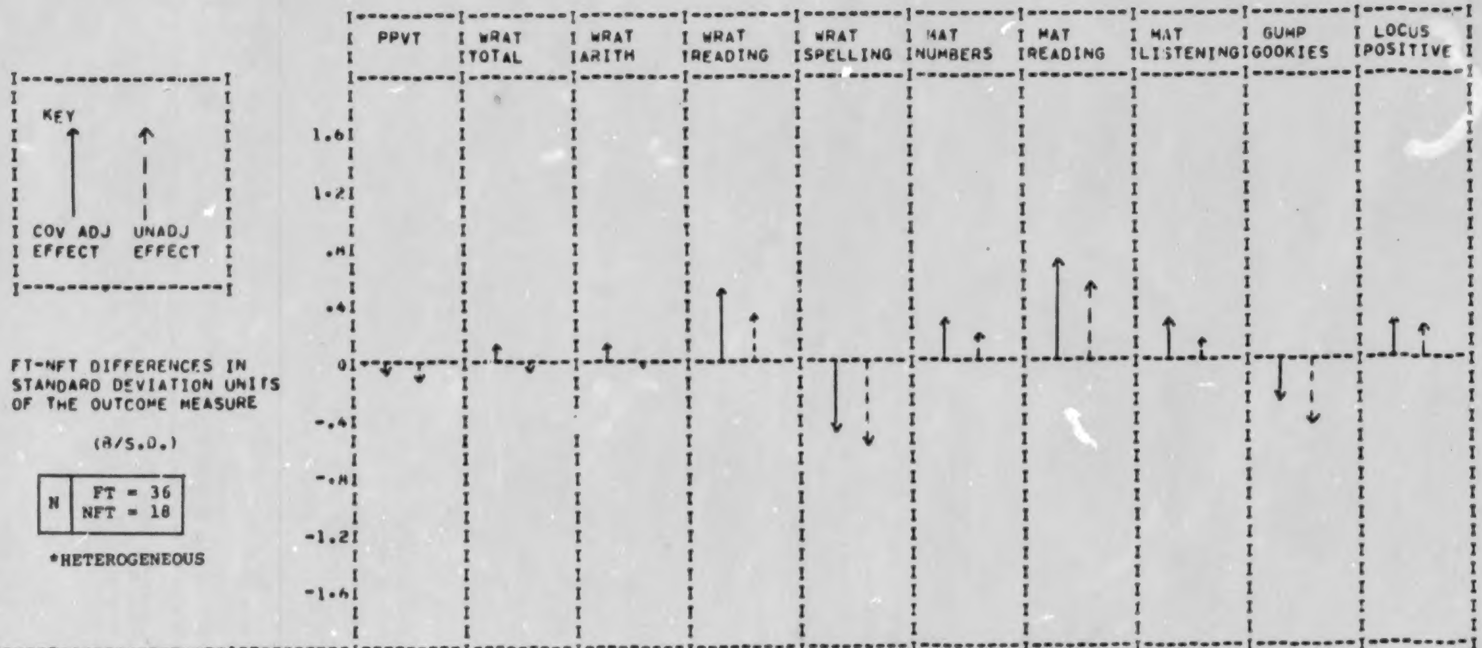
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.975	4.639	.286	1.508	2.844	1.619	-.256	-3.772	.031	-.217
	UNADJ	-1.128	2.437	-.118	.622	1.933	.847	-.776	-4.230	.002	-.151
S.E. = STANDARD ERROR OF B	ADJ	2.313	2.234	.860	1.189	1.264	1.388	1.118	1.469	.050	.499
	UNADJ	2.418	2.856	1.072	1.382	1.241	1.438	1.174	1.545	.048	.458
F= SIGNIFICANCE STATISTIC	ADJ	.178	4.296	.111	1.610	5.064	1.361	.053	6.587	.367	.189
	UNADJ	.217	.728	.012	.202	2.426	.347	.437	7.496	.002	.109
SD=RAW SCORE STANDARD DEVIATION		8.118	9.636	3.592	4.639	4.256	4.834	3.950	5.550	.159	1.522
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.120	.481	.080	.325	.668	.335	-.065	-.680	.192	-.142
	UNADJ	-.139	.253	-.033	.134	.454	.175	-.197	-.762	.013	-.099

FIGURE C-20 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE SITE NE



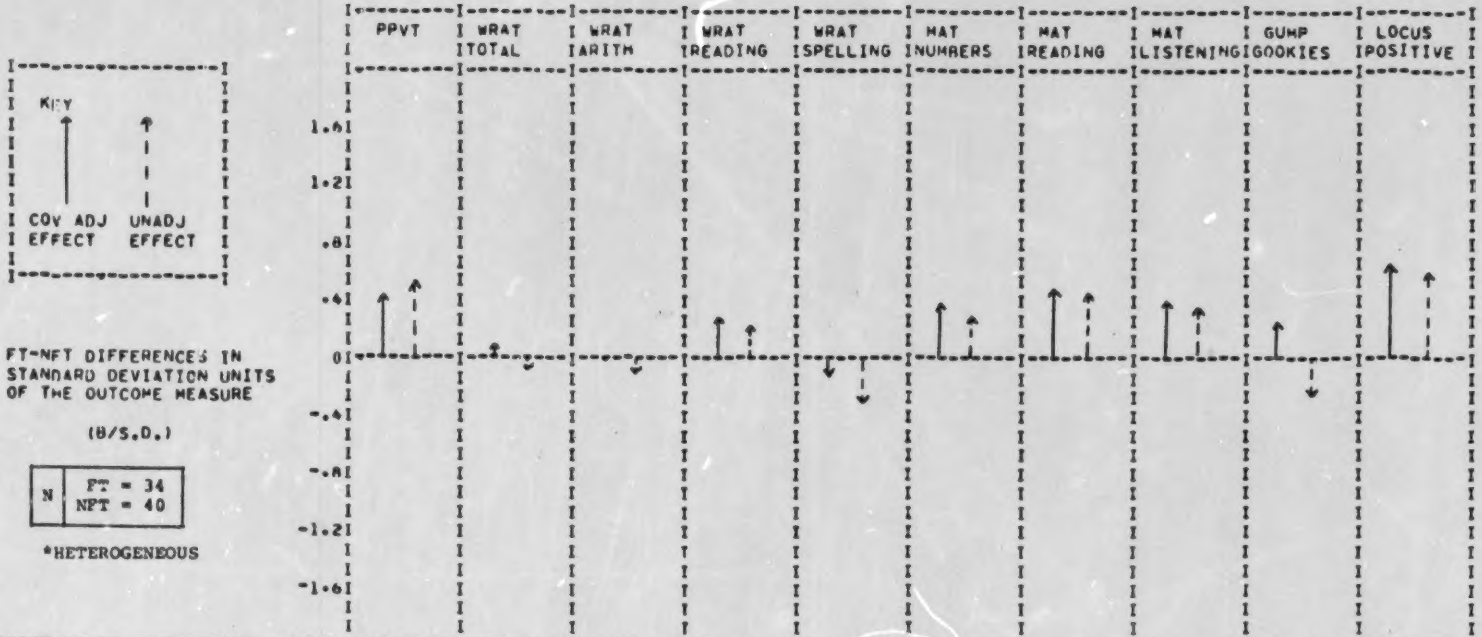
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
H= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.733	.048	-.972	.200	.820	-1.518	-1.229	-1.736	-.025	.621
	UNADJ	-.208	.641	-.760	.339	1.061	-1.292	-1.012	-1.559	-.023	.627
S.E. = STANDARD ERROR OF H	ADJ	1.829	1.635	.675	.837	.879	.738	.745	1.568	.039	.492
	UNADJ	2.274	2.510	.868	1.134	1.050	.896	.906	1.607	.040	.492
F= SIGNIFICANCE STATISTIC	ADJ	.161	.001	2.074	.057	.870	4.231	2.721	1.226	.411	1.591
	UNADJ	.008	.065	.767	.089	1.021	2.079	1.248	.941	.331	1.624
SD=RAW SCORE STANDARD DEVIATION		9.536	10.528	3.660	4.757	4.434	3.812	3.834	6.782	.168	2.099
H/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.077	.005	-.266	.042	.185	-.398	-.321	-.256	-.149	.296
	UNADJ	-.022	.061	-.208	.071	.239	-.339	-.264	-.230	-.137	.299

FIGURE C-21 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE SITE NC



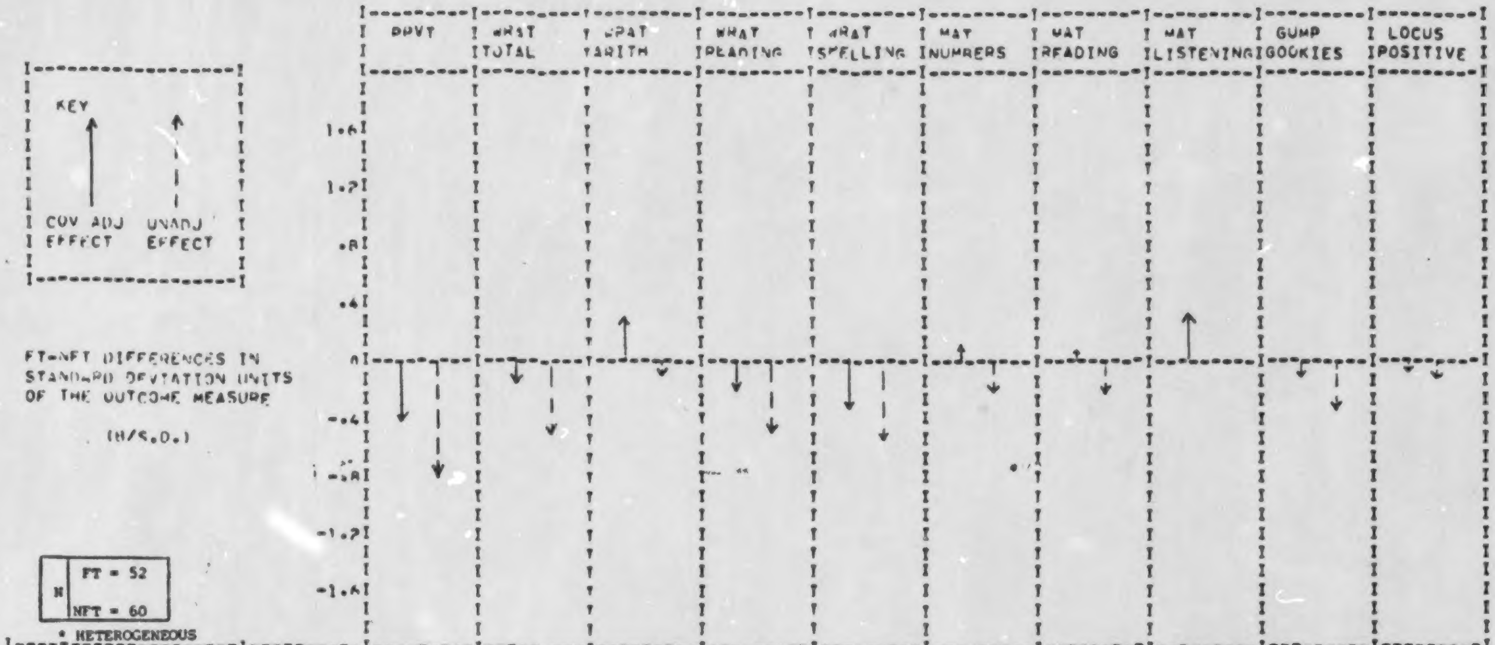
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.341	1.071	.458	2.469	-1.857	1.697	3.400	1.542	-.048	.440
	UNADJ	-.722	-.556	-.029	1.667	-2.194	.972	2.639	.667	-.076	.361
S.E. = STANDARD ERROR OF B	ADJ	.991	1.846	.799	.989	.937	1.504	1.097	1.343	.043	.490
	UNADJ	1.713	2.810	.989	1.403	1.069	1.656	1.371	1.661	.048	.491
F= SIGNIFICANCE STATISTIC	ADJ	.119	.351	.329	6.231	3.926	1.273	9.610	1.317	1.274	.804
	UNADJ	.178	.039	.001	1.412	4.217	.345	3.705	.161	2.583	.540
SD=RAW SCORE STANDARD DEVIATION		5.887	9.644	3.394	4.878	3.813	5.701	4.869	5.709	.168	1.694
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.058	.111	.135	.506	-.487	.298	.698	.270	-.288	.259
	UNADJ	-.123	-.048	-.008	.342	-.576	.171	.542	.117	-.457	.213

FIGURE C -22 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE SITE W1



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	* 3.874	* .517	-.021	* 1.069	-.532	1.623	2.148	2.395	.040	1.196
	UNADJ	4.913	-.76[illegible]	-.325	.877	-1.316	1.252	2.010	2.046	-.042	1.058
S.E. = STANDARD ERROR OF B	ADJ	1.889	1.565	.703	.831	.802	.793	.819	1.165	.039	.402
	UNADJ	2.121	2.108	.775	.973	.941	1.012	1.051	1.388	.039	.419
F= SIGNIFICANCE STATISTIC	ADJ	4.207	.109	.001	1.654	.440	4.192	6.886	4.228	1.079	8.684
	UNADJ	5.366	.132	.176	.811	1.958	1.529	3.655	2.172	1.200	6.748
SD=RAW SCORE STANDARD DEVIATION		9.361	8.983	3.304	4.168	4.059	4.354	4.589	5.999	.167	1.865
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.414	.058	-.006	.257	-.131	.373	.468	.399	.241	.636
	UNADJ	.525	-.085	-.099	.210	-.324	.287	.438	.341	-.255	.583

FIGURE C-23 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE SITE W2

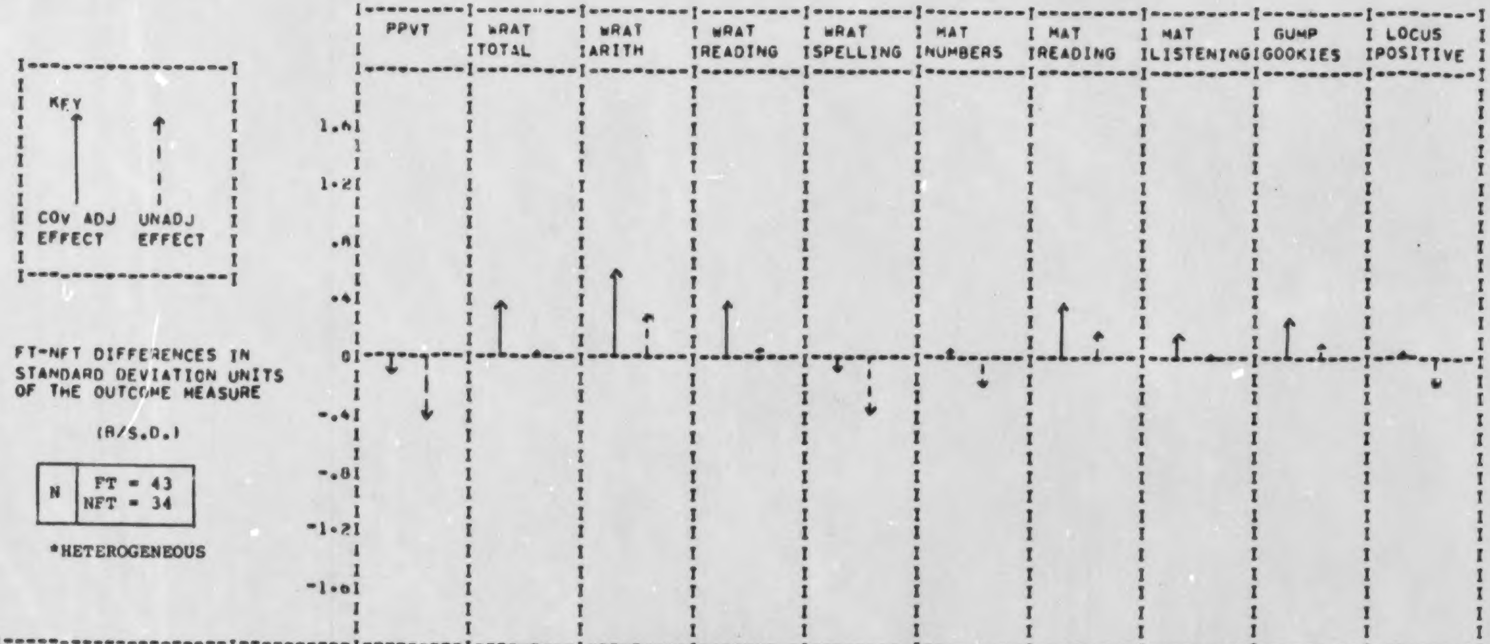


N FT = 52
NFT = 60

* HETEROGENEOUS

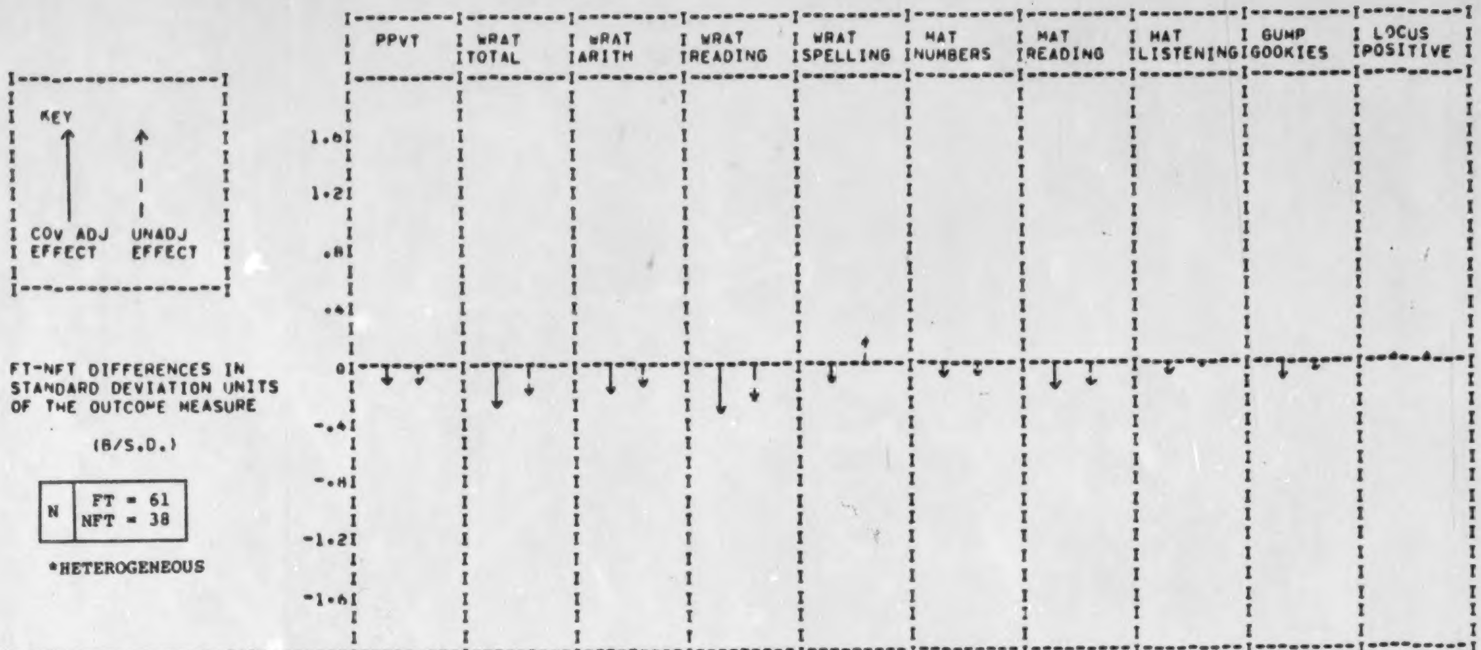
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-2.331	-1.512	1.002	-1.123*	-1.390	.389	.298	2.052	-.015	-.116
	UNADJ	-4.676	-5.429	-.332	-2.750	-2.347	-.649	-1.157	.000	-.049	-.202
S.E. = STANDARD ERROR OF B	ADJ	1.222	1.656	.630	1.071	.946	.674	1.114	1.355	.032	.427
	UNADJ	1.214	2.416	.882	1.292	.977	.761	1.154	1.442	.032	.394
F= SIGNIFICANCE STATISTIC	ADJ	3.637	.834	2.530	1.100	2.159	.333	.072	2.293	.220	.074
	UNADJ	14.836	5.050	.142	4.530	5.771	.727	1.005	.000	2.345	.263
SD=RAW SCORE STANDARD DEVIATION		5.552	10.374	3.409	5.527	4.217	3.169	4.814	5.975	.138	1.647
B/SD = EFFECT IN SD UNITS (SFE GRAPH)	ADJ	-.420	-.146	.294	-.203	-.330	.123	.062	.343	-.109	-.070
	UNADJ	-.842	-.523	.097	-.497	-.556	-.205	-.240	.000	-.350	-.122

FIGURE C-24 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE SITE S



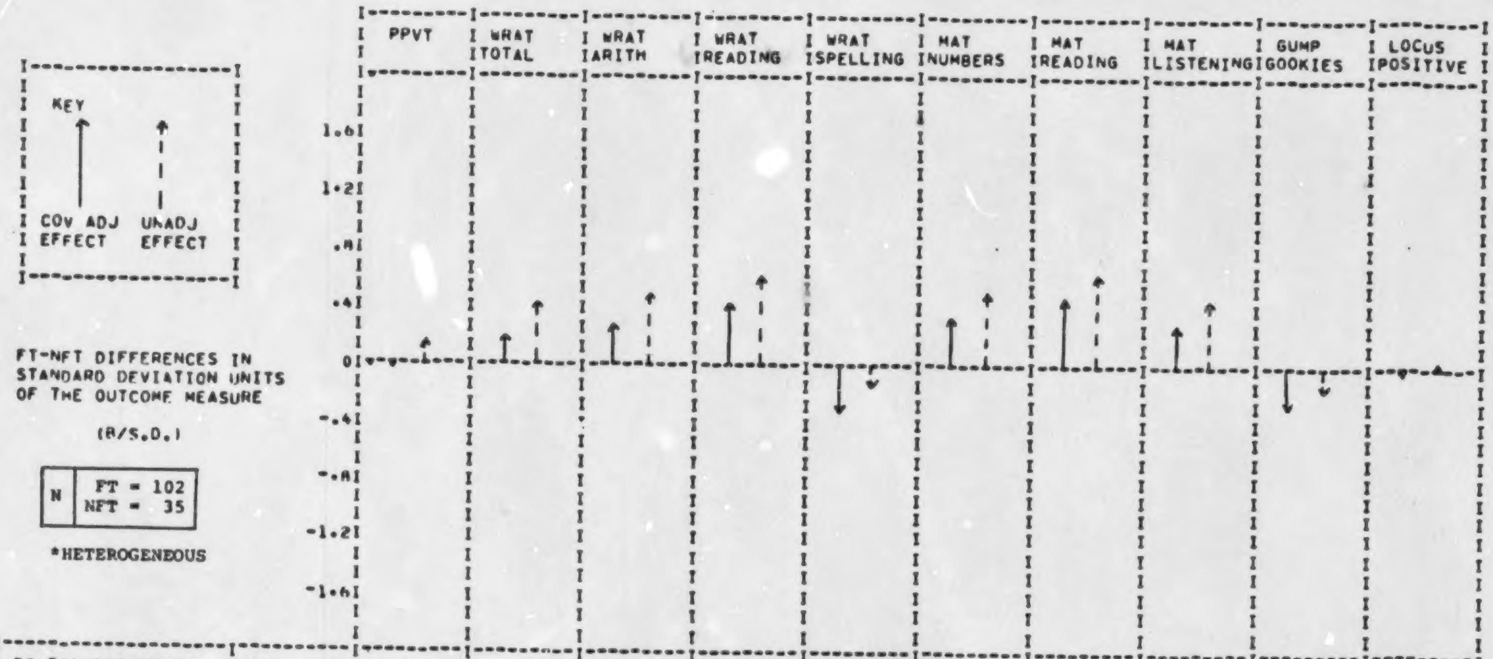
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	* -.979	4.083	* 2.522	* 1.890	-.328	* .294	1.927	1.306	.044	.051
	UNADJ	-3.608	.133	1.226	.273	-1.366	-1.170	.994	.014	.016	-.364
S.E. = STANDARD ERROR OF B	ADJ	1.413	1.660	.756	.866	.655	1.197	1.119	1.454	.036	.458
	UNADJ	1.891	2.506	.970	1.163	.791	1.416	1.197	1.629	.037	.478
F= SIGNIFICANCE STATISTIC	ADJ	.479	6.061	11.121	4.760	.251	.060	2.965	.807	1.473	.012
	UNADJ	3.641	.093	1.598	.055	2.979	.682	.690	.000	.186	.578
SD=RAW SCORE STANDARD DEVIATION		8.391	10.864	4.249	5.044	3.496	6.167	5.212	7.062	.161	2.081
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.117	.376	.593	.375	-.094	.048	.370	.185	.275	.024
	UNADJ	-.430	.012	.289	.054	-.391	-.190	.191	.002	.099	-.175

FIGURE C-25 KINDERGARTEN EFFECTS PROFILE FOR HIGH SCOPE SITE W3



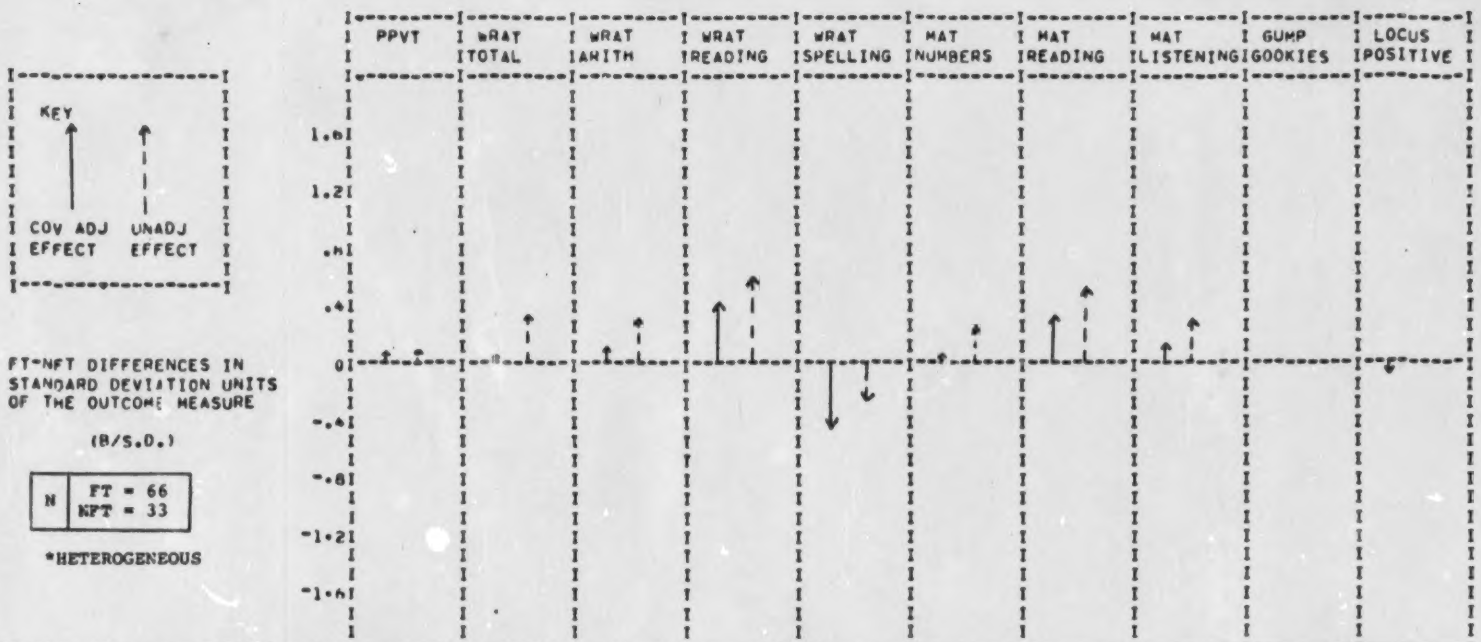
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-1.059	-2.981	-.684	-1.742	-.555	-.507	-.740	-.446	-.017	.011
	UNADJ	-.933	-2.115	-.486	-1.273	-.256	-.329	-.617	-.039	-.009	.053
S.E. = STANDARD ERROR OF B	ADJ	1.370	1.562	.632	.819	.795	.910	.767	1.164	.030	.333
	UNADJ	1.680	2.157	.718	1.042	.926	1.020	.859	1.310	.031	.339
F= SIGNIFICANCE STATISTIC	ADJ	.598	3.642	1.174	4.519	.657	.310	.933	.147	.321	.001
	UNADJ	.308	.961	.458	1.494	.736	.104	.517	.001	.079	.025
SD=RAW SCORE STANDARD DEVIATION		8.102	10.435	3.464	5.060	4.460	4.913	4.145	6.306	.149	1.631
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.131	-.286	-.198	-.344	-.124	-.103	-.179	-.071	-.114	.007
	UNADJ	-.115	-.203	-.140	-.252	.178	-.067	-.149	-.006	-.058	.033

FIGURE C-26 KINDERGARTEN EFFECTS PROFILE FOR FLORIDA SITE S1



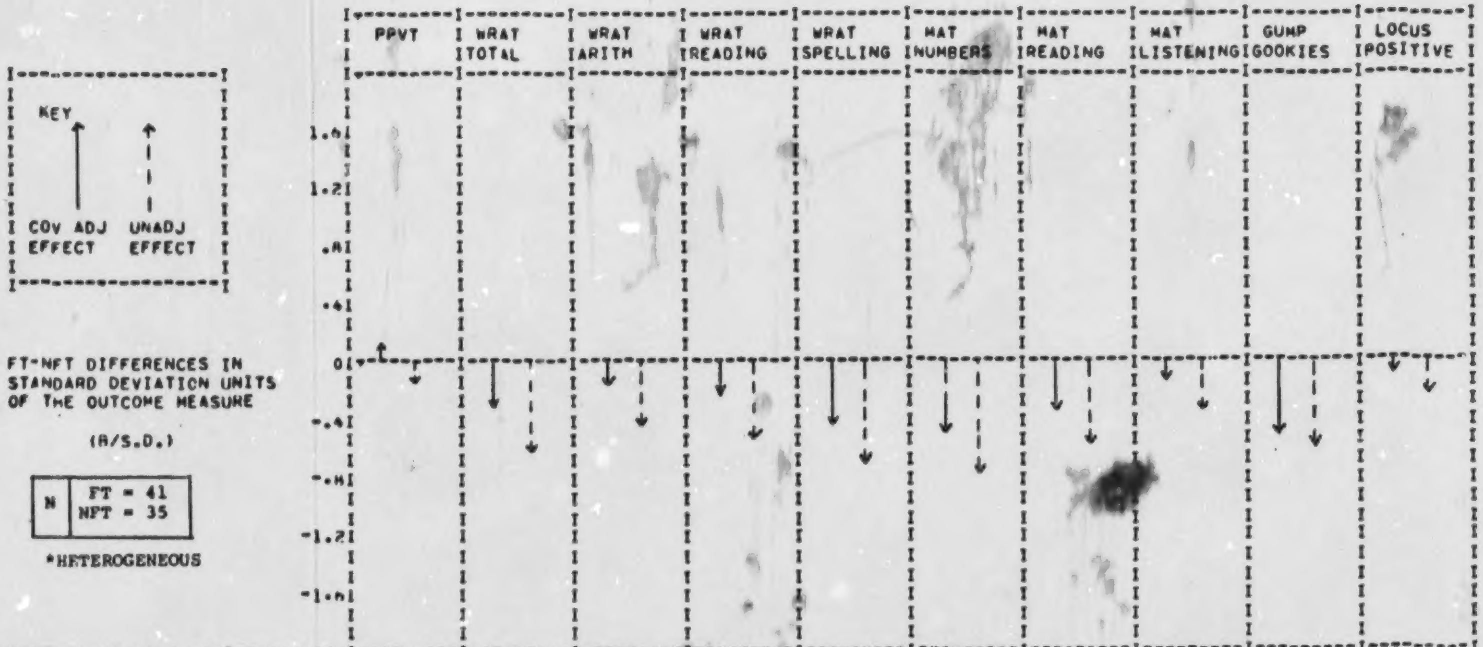
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.259	1.913	1.024	2.275	-1.386	1.744	2.244	1.622	-.041	-.057
	UNADJ	1.218	4.609	1.850	3.336	-.576	2.575	3.013	2.734	-.024	.046
S.E. = STANDARD ERROR OF B	ADJ	1.236	1.685	.640	.913	.799	.922	.809	1.051	.028	.342
	UNADJ	1.803	2.173	.755	1.051	.886	1.009	.891	1.156	.028	.349
F= SIGNIFICANCE STATISTIC	ADJ	.044	1.289	2.557	6.207	3.012	3.581	7.697	2.382	2.049	.028
	UNADJ	.456	4.498	6.006	10.074	.422	6.516	11.424	5.595	.686	.018
SD=RAW SCORE STANDARD DEVIATION		9.083	11.115	3.881	5.484	4.463	5.196	4.673	5.935	.145	1.767
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.028	.172	.264	.415	-.311	.336	.480	.273	-.282	-.032
	UNADJ	.134	.415	.477	.608	-.129	.495	.645	.461	-.163	.026

FIGURE C -27 KINDERGARTEN EFFECTS PROFILE FOR FLORIDA SITE S2



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	.421	.538	.376	*2.062	-1.900	.252	1.580	*.781	MISSING	* -.137
	UNADJ	.519	3.156	1.139	3.073	-1.055	1.302	2.701	1.998	MISSING	.000
S.E. = STANDARD ERROR OF B	ADJ	1.010	1.459	.688	.800	.810	.980	.779	1.252		.477
	UNADJ	1.393	2.182	.796	1.043	.912	1.143	1.088	1.485		.496
F= SIGNIFICANCE STATISTIC	ADJ	.173	.135	.298	6.635	5.503	.066	4.109	.389		.082
	UNADJ	.139	2.092	2.048	8.047	1.339	1.296	6.159	1.810		.000
SD=RAW SCORE STANDARD DEVIATION		6.319	10.003	3.646	5.118	4.161	5.218	5.094	6.798		1.850
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.067	.054	.103	.403	-.456	.048	.310	.115		-.074
	UNADJ	.082	.316	.312	.600	-.253	.249	.530	.294		.000

FIGURE C-28 KINDERGARTEN EFFECTS PROFILE FOR FLORIDA SITE W

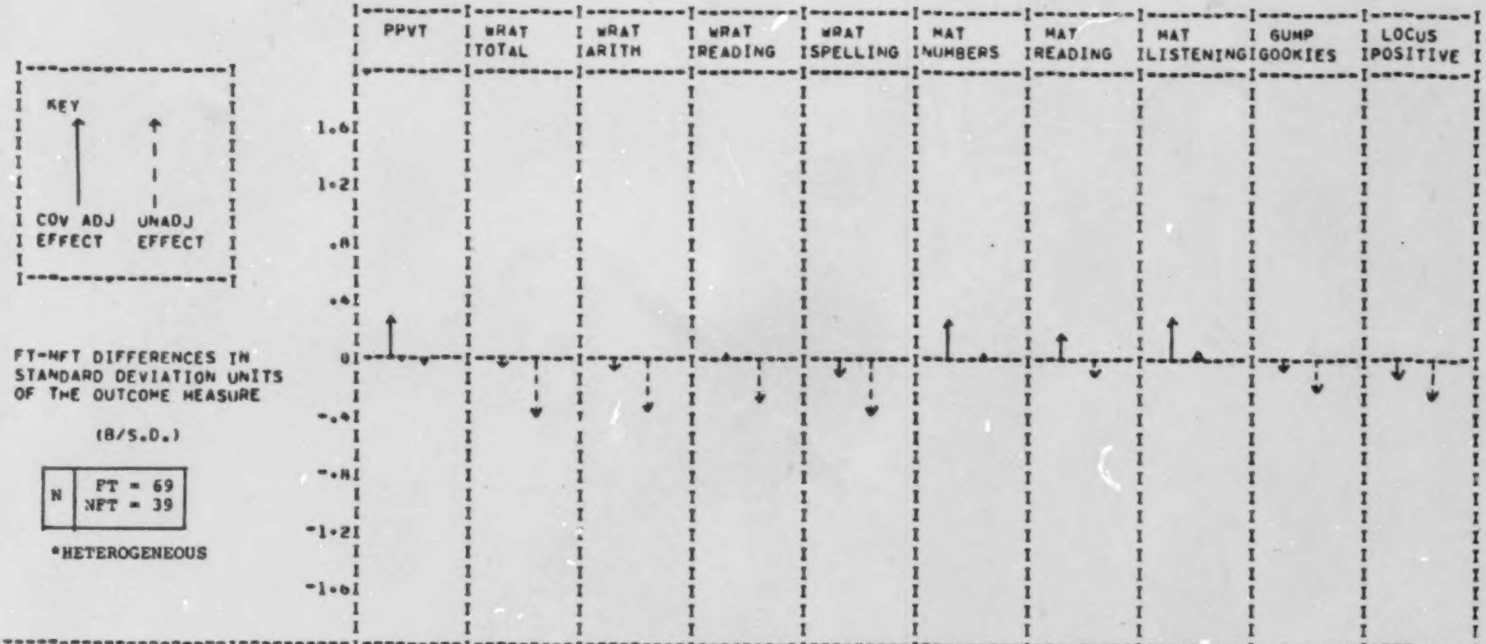


N	FT = 41 NFT = 35

*HETEROGENEOUS

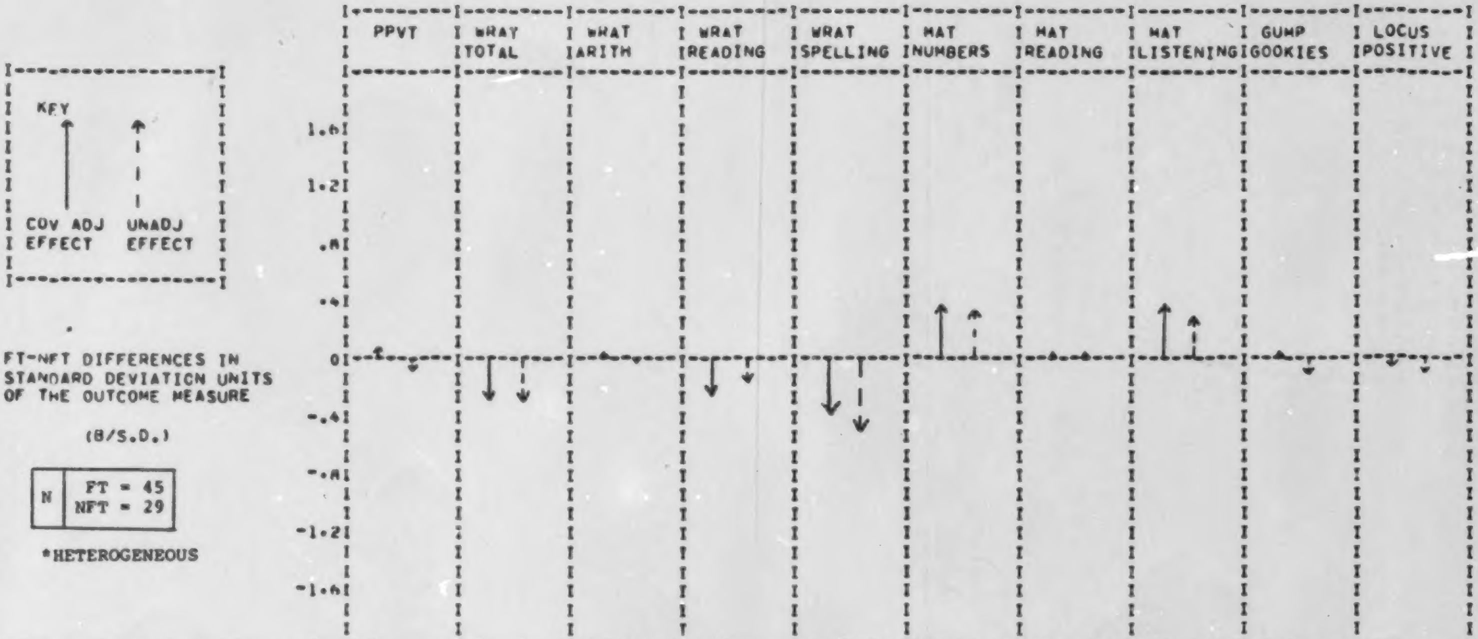
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	* .557	-4.278	-.602	-1.452	-2.224	* -3.179	-1.758	-.766	-.079	-.140
	UNADJ	-.823	-8.459	-1.587	-3.438	-3.434	-4.944	-2.891	-2.002	-.093	-.453
S.E. = STANDARD ERROR OF B	ADJ	.890	1.781	.660	.983	.891	.930	.956	1.236	.035	.463
	UNADJ	1.295	2.899	.841	1.465	1.092	1.361	1.120	1.406	.035	.484
F= SIGNIFICANCE STATISTIC	ADJ	.392	5.770	.831	2.182	6.226	11.680	3.386	.384	5.088	.091
	UNADJ	.404	8.516	3.559	5.506	9.895	13.198	6.658	2.029	7.035	.875
SD=RAW SCORE STANDARD DEVIATION		5.606	13.212	3.719	6.554	5.017	6.376	5.049	6.150	.158	2.102
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.099	-.324	-.162	-.222	-.443	-.499	-.348	-.125	-.496	-.067
	UNADJ	-.147	-.640	-.427	-.524	-.684	-.775	-.573	-.326	-.588	-.215

FIGURE C 29 KINDERGARTEN EFFECTS PROFILE FOR FLORIDA SITE NC



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
H= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	1.675	-.567	-.222	.180	-.526	1.492	.720	1.854	-.010	-.230
	UNADJ	-.232	-3.763	-1.010	-1.130	-1.622	.149	-.517	.178	-.030	-.463
S.E. = STANDARD ERROR OF B	ADJ	.915	1.442	.495	.761	.756	.914	.843	1.276	.029	.361
	UNADJ	1.177	1.872	.576	.879	.844	1.075	.924	1.386	.028	.352
F= SIGNIFICANCE STATISTIC	ADJ	3.351	.155	.201	.056	.484	2.655	.729	2.111	.127	.407
	UNADJ	.039	4.040	3.075	1.655	3.690	.019	.314	.017	1.138	1.725
SD=RAW SCORE STANDARD DEVIATION		5.847	9.476	2.903	4.400	4.268	5.341	4.595	6.887	.142	1.765
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.286	-.060	-.076	.041	-.123	.279	.157	.269	-.073	-.130
	UNADJ	-.040	-.397	-.348	-.257	-.380	.028	-.113	.026	-.214	-.262

FIGURE C-30 KINDERGARTEN EFFECTS PROFILE FOR EDC SITE NE1

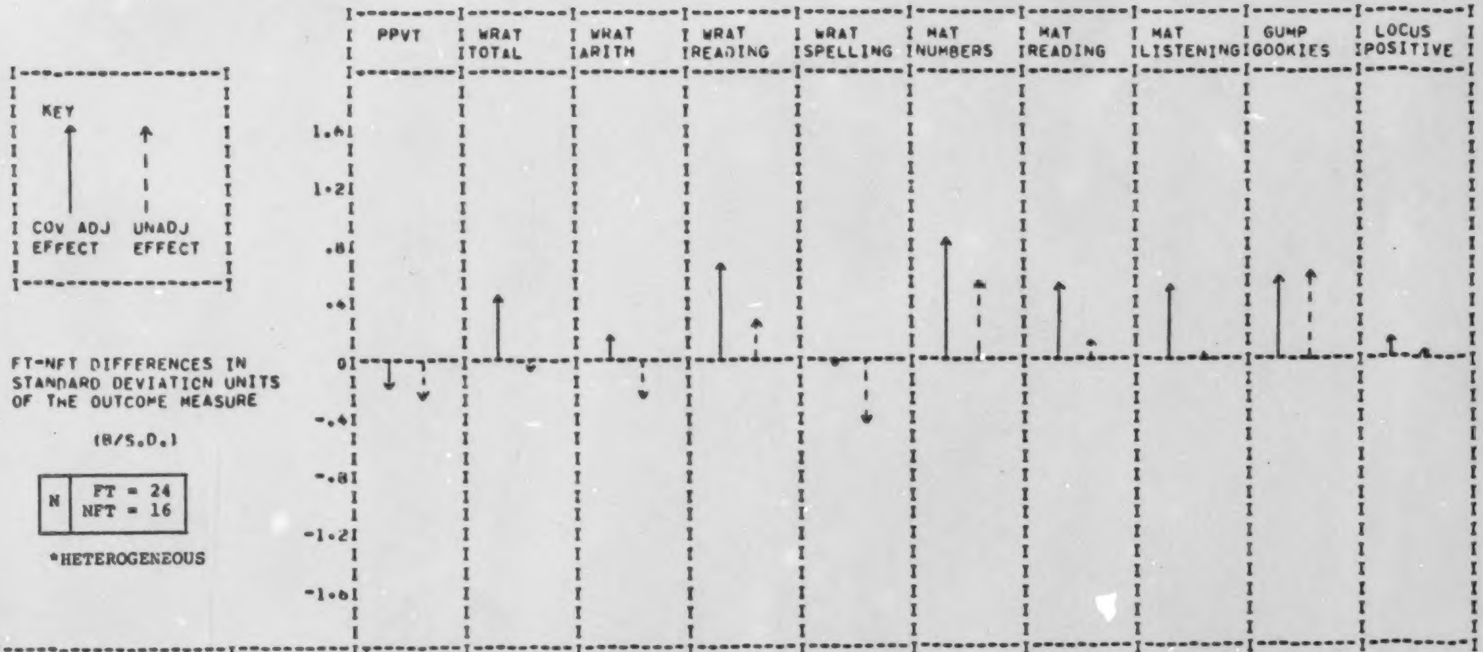


N	
FT = 45	NFT = 29

*HETEROGENEOUS

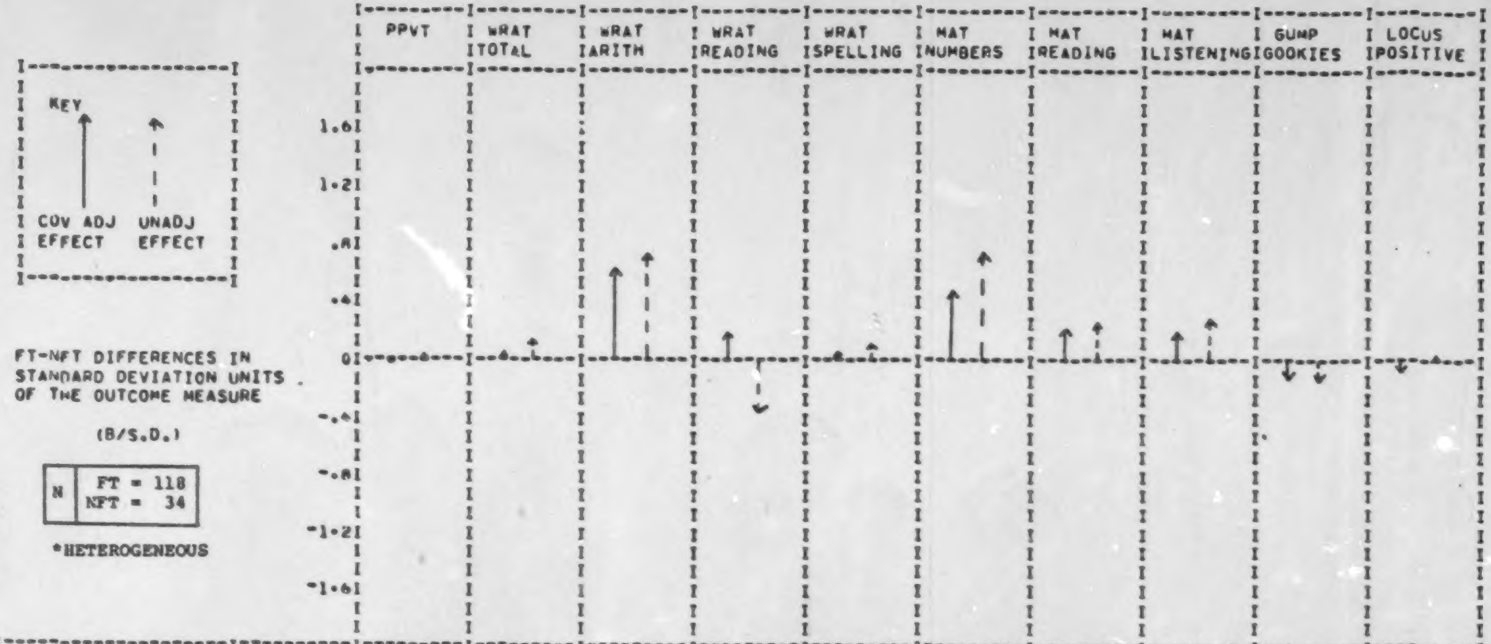
		PPVT	WRAY TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	* .535	-3.006	.080	-1.361	-1.726	1.492	.124	2.640	.004	-.124
	UNADJ	-.841	-3.140	-.025	-.910	-2.205	1.315	.217	2.060	-.013	-.202
S.E. = STANDARD ERROR OF B	ADJ	1.772	2.090	.660	1.431	.854	.893	1.043	1.520	.034	.522
	UNADJ	2.299	2.201	.782	1.560	1.020	.958	1.172	1.642	.034	.505
F= SIGNIFICANCE STATISTIC	ADJ	.091	2.069	.015	.904	4.085	2.790	.014	3.018	.015	.056
	UNADJ	.134	2.036	.001	.341	4.674	1.885	.034	1.574	.138	.159
SD=RAW SCORE STANDARD DEVIATION		9.599	11.281	3.261	6.520	4.389	4.046	4.889	6.921	.143	2.094
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.056	-.267	.025	- .242	- .393	.369	.025	.382	.029	-.059
	UNADJ	-.088	-.278	-.008	- .139	- .502	.325	.044	.298	-.090	-.096

FIGURE C-31 KINDERGARTEN EFFECTS PROFILE FOR EDC SITE NE2



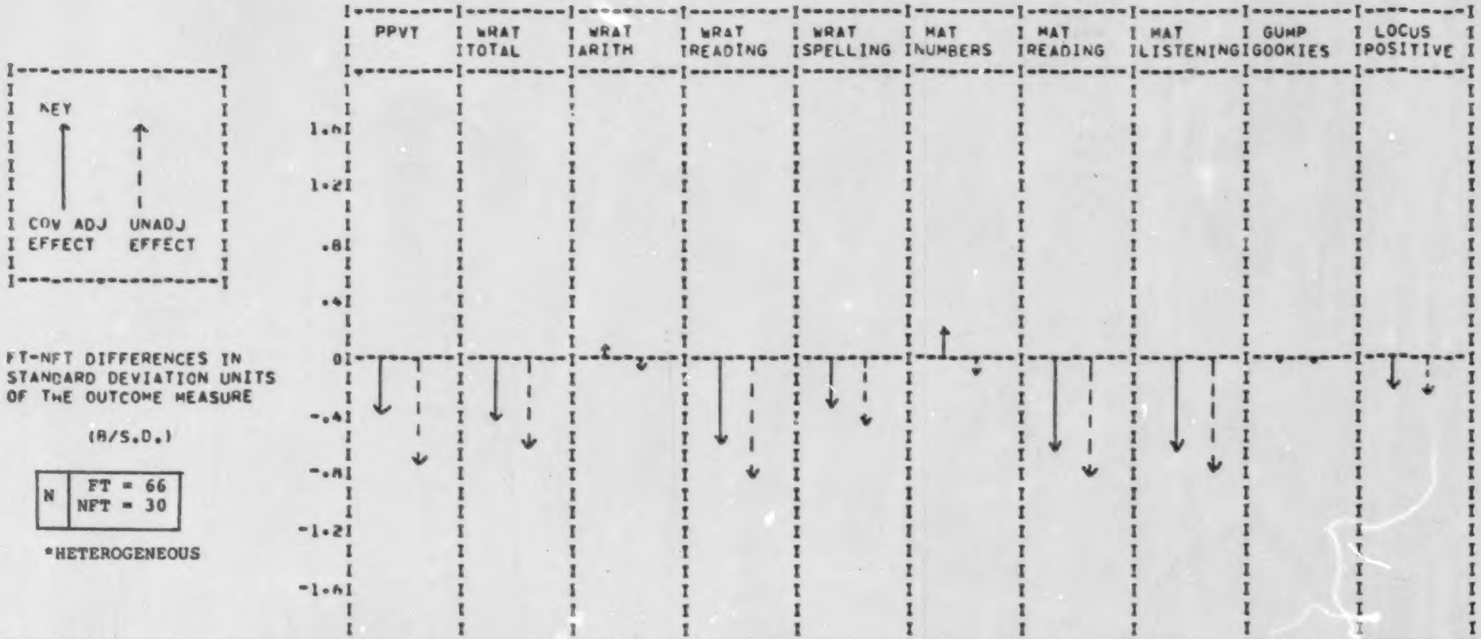
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	* -2.234	5.058	.579	4.661	-.122	4.197	2.239	3.403	.080	.250
	UNADJ	-2.945	-.577	-.918	1.967	-1.626	2.775	.489	.225	.085	.071
S.E. = STANDARD ERROR OF B	ADJ	3.500	2.811	1.166	1.954	1.155	1.639	1.407	1.839	.052	.626
	UNADJ	3.895	3.873	1.245	2.337	1.307	1.623	1.483	2.180	.047	.595
F= SIGNIFICANCE STATISTIC	ADJ	.407	3.238	.247	5.544	.011	6.556	2.533	3.426	2.316	.159
	UNADJ	.572	.027	.543	.709	1.549	2.921	.109	.011	3.274	.014
SD=RAW SCORE STANDARD DEVIATION		11.684	11.535	3.735	7.022	3.969	5.016	4.423	6.492	.143	1.724
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.191	.438	.155	.655	-.031	.837	.506	.524	.557	.145
	UNADJ	-.252	-.050	-.246	.280	-.410	.553	.111	.035	.598	.041

FIGURE C-32 KINDERGARTEN EFFECTS PROFILE FOR EDC SITE S



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP COOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-.010	.198	1.846	7.815	.167	5.076	.978	1.659	-.017	-.137
	UNADJ	.186	1.076	2.123	-1.464	.419	5.633	1.306	2.004	-.019	.038
S.E. = STANDARD ERROR OF B	ADJ	.674	1.070	.448	.578	.591	1.112	.762	1.114	.031	.337
	UNADJ	.893	1.661	.546	.782	.724	1.344	.921	1.328	.031	.349
F= SIGNIFICANCE STATISTIC	ADJ	.000	.034	16.983	182.476	.079	20.834	1.646	1.717	.308	.100
	UNADJ	.043	.422	15.109	3.502	.335	17.569	2.010	2.275	.385	.012
SD=RAW SCORE STANDARD DEVIATION		4.573	8.516	2.934	4.051	3.709	7.273	4.749	6.854	.159	1.789
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.002	.023	.629	1.929	.045	.698	.206	.213	-.108	-.060
	UNADJ	.041	.127	.724	-.361	.113	.774	.275	.292	-.121	.021

FIGURE C-33 KINDERGARTEN EFFECTS PROFILE FOR PITTSBURGH SITE NC



FT-NFT DIFFERENCES IN STANDARD DEVIATION UNITS OF THE OUTCOME MEASURE

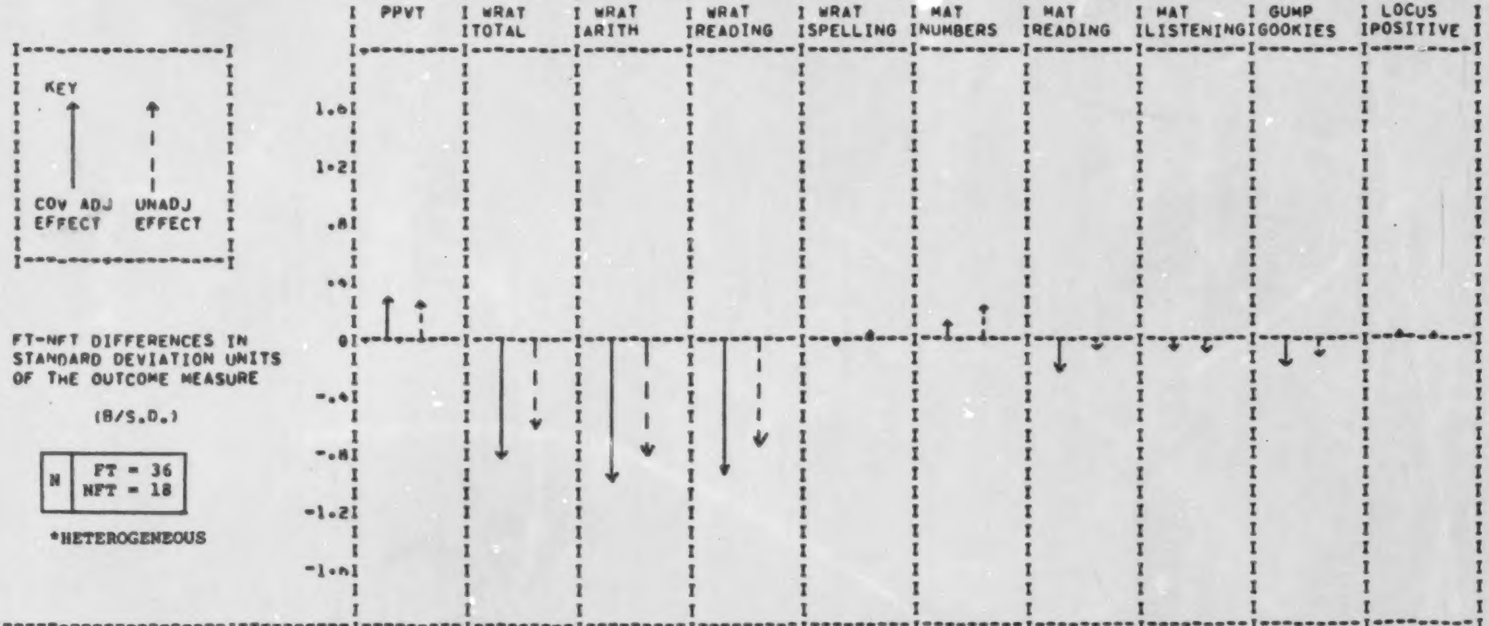
(B/S.D.)

N	FT = 66 NFT = 30

*HETEROGENEOUS

		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-2.378	-4.391	.234	-3.076	-1.549	1.063	-3.369	-4.315	-.003	-.403
	UNADJ	-4.530	-6.570	-.230	-4.261	-2.079	-.594	-4.282	-5.179	-.002	-.427
S.E. = STANDARD ERROR OF B	ADJ	1.041	1.412	.655	.840	.867	.955	.923	1.265	.036	.444
	UNADJ	1.284	2.193	.784	1.056	.967	1.189	1.082	1.378	.035	.437
F= SIGNIFICANCE STATISTIC	ADJ	5.219	9.668	.127	13.419	3.187	1.238	13.330	11.626	.007	.822
	UNADJ	12.453	8.975	.086	16.288	4.618	.249	15.669	14.128	.005	.958
SD=RAW SCORE STANDARD DEVIATION		6.172	10.369	3.542	5.166	4.476	5.379	5.278	6.676	.159	1.982
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.385	-.423	.066	-.595	-.346	.198	-.638	-.646	-.020	-.203
	UNADJ	-.734	-.634	-.065	-.825	-.464	-.110	-.811	-.776	-.015	-.216

FIGURE C-34 KINDERGARTEN EFFECTS PROFILE FOR PITTSBURGH SITE NE

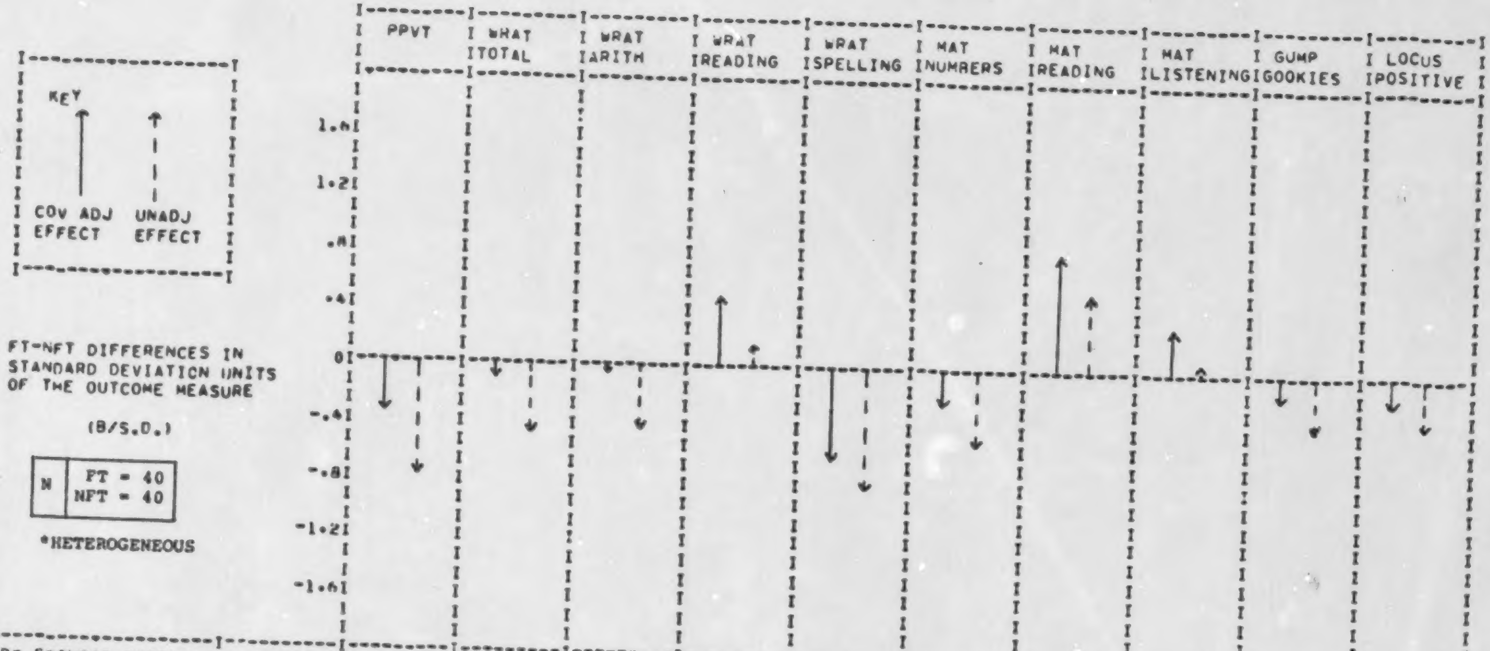


N	FT = 36 NFT = 18

*HETEROGENEOUS

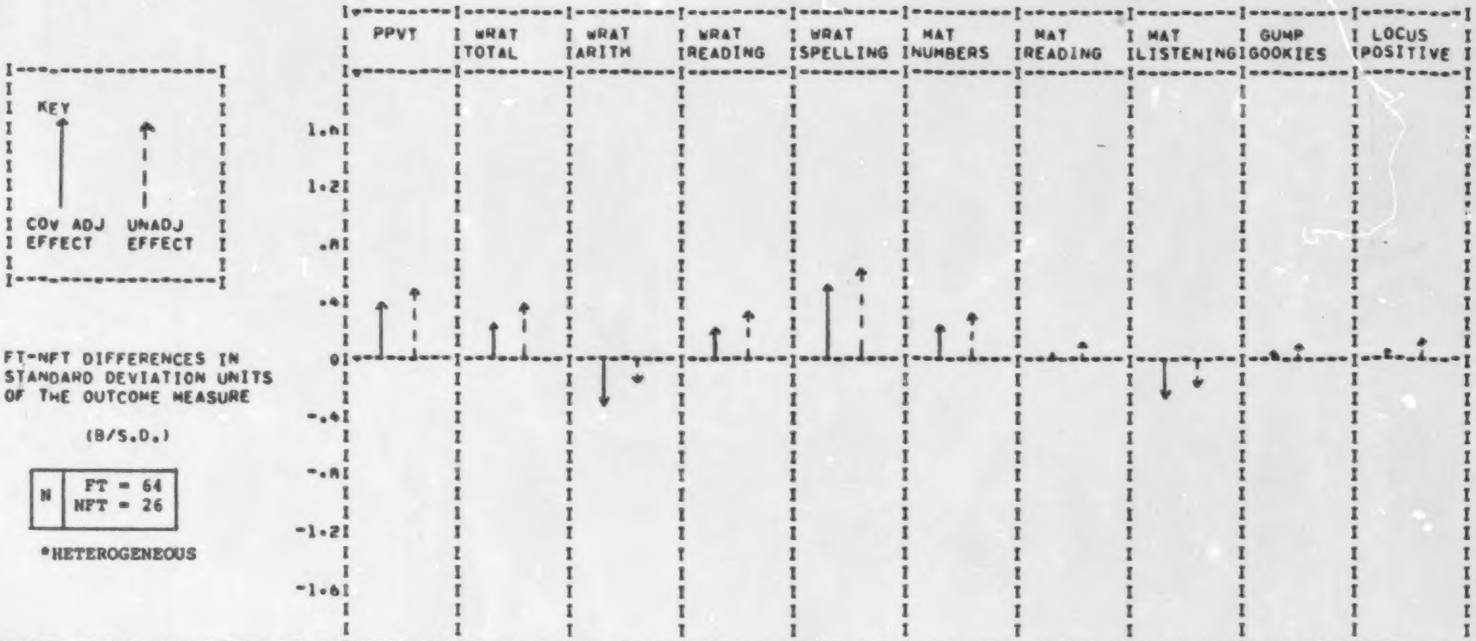
		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	3.039	-8.714	-4.268	-4.326	-.120	.469	-1.101	-.217	-.023	.046
	UNADJ	2.835	-6.615	-3.497	-3.345	.227	.806	-.290	-.309	-.015	.007
S.E. = STANDARD ERROR OF B	ADJ	2.880	2.661	1.145	1.186	1.252	1.130	1.337	1.231	.038	.692
	UNADJ	3.303	3.146	1.236	1.358	1.294	1.116	1.448	1.206	.037	.656
F= SIGNIFICANCE STATISTIC	ADJ	1.113	10.722	13.888	13.304	.009	.172	.679	.031	.357	.004
	UNADJ	.737	4.535	8.005	6.066	.031	.521	.040	.066	.156	.000
SD=RAW SCORE STANDARD DEVIATION		10.951	10.669	4.376	4.730	4.262	3.692	4.769	3.973	.122	2.139
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.277	-.817	-.975	-.915	-.028	.127	-.231	-.055	-.186	.022
	UNADJ	.259	-.620	-.799	-.707	.053	.218	-.061	-.078	-.121	.003

FIGURE C-35 KINDERGARTEN EFFECTS PROFILE FOR SEDL SITE ME



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	-3.491	-1.093	-.170	2.448	-3.371	-.999	2.840	1.473	-.025	-.449
	UNADJ	-7.725	-6.350	-1.650	.650	-4.350	-2.100	1.900	.300	-.054	-.925
S.E. = STANDARD ERROR OF B	ADJ	1.679	1.809	.630	.882	1.030	.823	.705	.989	.033	.615
	UNADJ	2.066	2.429	.848	1.105	1.056	.883	.751	1.058	.034	.604
F= SIGNIFICANCE STATISTIC	ADJ	4.324	.365	.073	7.703	10.709	1.472	16.227	2.219	.564	.533
	UNADJ	13.984	4.852	3.788	.346	16.969	5.651	6.394	.080	2.536	2.349
SD=RAW SCORE STANDARD DEVIATION		9.969	11.123	3.858	4.921	5.178	4.065	3.473	4.704	.153	2.705
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	-.350	-.098	-.044	.498	-.651	-.246	.818	.313	-.161	-.166
	UNADJ	-.775	-.481	-.428	.132	-.840	-.517	.547	.064	-.355	-.342

FIGURE C-36 KINDERGARTEN EFFECTS PROFILE FOR SEDL SITE W



		PPVT	WRAT TOTAL	WRAT ARITH	WRAT READING	WRAT SPELLING	MAT NUMBERS	MAT READING	MAT LISTENING	GUMP GOOKIES	LOCUS POSITIVE
B= SPONSOR EFFECT (FT-NFT DIFFERENCES)	ADJ	3.595*	2.689	-.956	1.399	2.247	.847	.046	-1.575	.007	.120
	UNADJ	4.582	4.214	-.496	2.075	2.635	1.297	.486	-1.026	.014	.354
S.E. = STANDARD ERROR OF B	ADJ	1.407	1.800	.558	1.197	.869	.748	.913	1.240	.034	.536
	UNADJ	2.082	2.400	.733	1.389	.956	.932	1.012	1.329	.034	.568
F= SIGNIFICANCE STATISTIC	ADJ	6.530	2.232	2.932	1.364	6.682	1.284	.002	1.612	.040	.050
	UNADJ	4.841	3.084	.458	2.230	7.603	1.939	.230	.596	.167	.387
SD=RAW SCORE STANDARD DEVIATION		9.191	10.495	3.161	6.049	4.277	4.050	4.360	5.736	.152	2.502
B/SD = EFFECT IN SD UNITS (SEE GRAPH)	ADJ	.391	.256	-.303	.231	.525	.209	.010	-.275	.045	.048
	UNADJ	.499	.462	-.157	.343	.616	.320	.111	-.179	.093	.141

FIGURE C-37 KINDERGARTEN EFFECTS PROFILE FOR SEDL SITE 6

APPENDIX D

DESCRIPTION OF THE FOLLOW THROUGH SPONSORS

APPENDIX D

DESCRIPTION OF THE FOLLOW THROUGH SPONSORS

1.0 INTRODUCTION

The essential element of the planned variation aspect of Follow Through is the implementation of a variety of educational approaches by program Sponsors. The Sponsor is the outside change agent responsible for working with individual projects to deliver a new approach to education in the project's classrooms. There are 22 Sponsors working with nearly 170 projects throughout the nation to develop and refine successful approaches to instruction. Although the instructional approaches vary, all Sponsors share common orientations.

- All of them seek to develop children's learning abilities.
- All recognize the importance of individual and small group instruction and frequent exchange between children and concerned adults.
- All are committed to making learning interesting and relevant to the child's cultural background.
- All believe that the child's success in learning is inseparable from his self-esteem, motivation, autonomy, and environmental support. (USOE, 1973)

While all Sponsors are committed to these orientations the degree of their commitment and their approach to operationalizing it varies widely. So too do the psychological and philosophical bases underlying each Sponsor's approach. Some are more oriented toward academic achievement while others are more concerned with developing a process of instruction which will instill a desire to learn. Still others are oriented to teaching how to learn. Some Sponsors appear to be very similar in approach, others widely diverse. Regardless of appearances, all Sponsors do differ; yet all pursue the educational and social objectives of the Follow Through program.

The concept of planned variation is intended to help determine which of a variety of possible educational approaches works best in which of a variety of settings. The program Sponsor is the basic building block of this effort.

Whereas there are 22 Sponsors in Project Follow Through, only ten of these have been included in the analyses and discussions of this Report. While these Sponsors are likely to be representative of the full spectrum of instructional approaches, they were chosen because they have sufficient information in the data base to make analysis of these data feasible. Each of these ten Sponsors is described below.

2.0 FAR WEST LABORATORY

Responsive Educational Program

Evolving from the belief that a healthy self-concept allows a child to appreciate himself, his culture, and both his abilities and limitations, this model provides the child a learning environment in which he can explore and discover. Within a carefully designed setting the student is free to choose those activities he wishes to engage in. The goal is for this freedom and exploration to result in the child making interrelated discoveries about his physical and social world, all the while developing a healthy self-concept, and knowledge.

This autotelic (self-revealing) approach holds that the best way for a child to learn is for him to be in an environment in which he can try things out, risk, guess, ask questions, and make discoveries without serious psychological consequences. The learning centers, tasks, and games utilized by this model structure such an environment to some extent. The materials and the child's interaction with them are self-rewarding and stimulate the development of self-direction and inner controls. Teachers provide guidance but the child works on his own. There is no set pace. Learning sequences have been developed but each student works at his own rate. The model assumes that no single theory of learning can account for all the modes in which children learn; therefore, it seeks to provide a variety of educational alternatives which build on the background, culture, and life-style the child brings with him to the classroom.

Objectives:

- Make available a variety of education alternatives in the classroom so that the child is free to explore and to set his own learning pace.
- Develop the instructional staff to become more responsive to the individual child's needs.
- Develop the problem solving abilities of the child.
- Help the child develop confidence in his own capacity to succeed.

- Help the child develop the academic skills necessary for effective problem solving.
- Develop a learning environment that helps the child make interrelated discoveries about his physical and social world and develop a healthy self-concept. (USOE, 1973)

3.0 UNIVERSITY OF ARIZONA

Tucson Early Education Model

This model holds that an educational program should provide a child with a variety of experiences which will develop both his academic and social ability to function effectively and confidently in society. The skills and abilities required for participation in contemporary society are missing in the behavioral repertoires of many individuals because their background does not provide an adequate basis for their development. The educational experiences of the Arizona model seek to overcome this perceived deficit. Skills are always taught in a functional setting, and concepts are illustrated with examples from areas both within and outside the classroom. Teachers individualize and emphasize adult-child interaction on a one-to-one basis. Recognizing the differences in needs and learning rates of children, a great variety of behavioral options of both a self-selected and structured nature are provided students.

The curriculum focuses on four general areas of development: language competence, development of an intellectual base, development of a motivational base, and social arts and skills. The classrooms in which these are elaborated are organized into behavioral settings and interest centers for small groups, to encourage interactions among the child, his environment, and others. In addition, this classroom organization encourages social reinforcement techniques while the curriculum materials used are in themselves arranged for their reinforcing value.

Objectives:

- Develop the child's ability to think as facilitating the learning process.
- Develop the child's social and academic skills toward effective social interaction and communication.
- Develop attitudes and behavioral patterns which will enhance the total learning and socialization process for the child. (USOE, 1973)

4.0 BANK STREET COLLEGE

Bank Street College of Education Approach

Bank Street believes that learning of specific skills should not take place independent of healthy emotional development. Learning and development are intertwined. Learning must be pursued by the child on behalf of his own development; if not, it will be superficial. Therefore, the Bank Street classroom is designed to offer a rational and democratic situation in which a child's positive image as a learner and a person can develop.

The classroom is the child's workroom. He participates actively in his own learning as the adults in his room support his autonomy while expanding his world and sensitizing him to the meanings of his experiences in it. Academic skills are acquired within the broad context of planned activities that provide appropriate ways of expressing and organizing children's interests. The classroom is a stable organized environment. The teacher introduces activities and plans events but always in terms of the individual child's response. The teaching is diagnostic with a strong emphasis on individualized follow-up. While the planned activities originate from classroom themes such as organizing chores, or block building, they later extend to community themes (marketing, traffic, and water safety).

In the Bank Street classroom the focus is on tasks that are satisfying in terms of the child's own goals and productive for his cognitive and affective development. Academic skills are learned in a context of a relevant, engaging classroom life.

Objectives:

- Provide an individualized curriculum.
- Enable children not only to acquire basic knowledge and skills but also to master how to learn.
- Encourage communication which is self-initiated, creative, and expressive.
- Develop agreed-upon limits for behavior with full freedom of expression within these limits.
- Create a learning environment to challenge the child and to stimulate and support probing and problem solving.
- Extend the learning experience beyond the walls of the classroom.

- Encourage reinforcing relationships within each teaching team and among all staff members.
- Involve parents in classrooms and in social and community activities related to the school.
- Provide opportunities for parent-staff interaction on behalf of the individual child and the total program. (USOE, 1973)

5.0 UNIVERSITY OF OREGON

University of Oregon Engleman/Becker Model for Direct Instruction

Engleman and Becker believe that children will learn if they are taught well and there is a payoff for learning. They insist that a child who fails is one who has not been taught properly and that the remedy lies in teaching the skills that have not been mastered. The model holds that disadvantaged children can perform at "normal" levels of achievement when the instructional program builds, at an accelerated pace, upon the skills they bring to school. Therefore, the primary concern of this compensatory program is to teach academic skills and teach them rapidly.

In the model's classrooms at least one hour a day is spent on academic skills--reading, arithmetic, and language--in small group situations. The use of reinforcement is a key element in this aspect of the program. Children are smiled at and praised for correct performance. The materials are programmed and sequenced so that the tasks a child encounters are not too difficult. The teacher works with only four to six children in a rapid paced question-answer model; the children respond in unison in a prescribed fashion. In this manner the teacher receives continuous feedback on the performance of children and children are immediately rewarded for good performance.

Objectives:

- Bring the child up to the normal level of achievement by building on the skills which the child brings to school.
- Achieve a faster-than-normal rate of mastery of basic learning skills. (USOE, 1973)

6.0 UNIVERSITY OF KANSAS

Behavior Analysis Approach

The behavior analysis model uses systematic reinforcement to encourage desired behavior. A token exchange system is set up to provide precise, positive reinforcement of desired behavior. The tokens provide an immediate reward for successful completion of a task. Later these may be exchanged for an activity that is desired, such as playing with blocks, listening to a story, or recess. The token system prevents the immediate delivery of a reinforcing activity from interfering with the behavior which is being rewarded.

The Kansas program holds that an effective system of reinforcement makes the reward contingent on improved academic or social performance. Yet the token system does not preclude the possibility that learning itself can be rewarding. The tokens are used only to support early efforts in a particular area. As the child achieves a level of mastery where the new skill itself is rewarding, the token reinforcement is decreased or discontinued. The teacher's role is that of a behavior modifier. She can monitor the child's progress by noting the amount of tokens he has available to exchange. Thus, the token system provides feedback to the teacher as well as the child.

The token system is used in both the social and academic areas. Children are reinforced for appropriate student role behavior, as well as for progress in reading, language, writing, and mathematics. Programmed instruction materials are used to allow for individualized instruction and to further facilitate teacher monitoring of rates of progress. The model calls for careful and accurate criteria and instructional objectives and this is made possible in large measure by the programmed instruction approach.

Objectives:

- Facilitate and accelerate the child's mastery of basic skills, particularly in reading and arithmetic, through the establishment of a "token economy" within classrooms.
- Train instructional staff to teach appropriate academic and social skills through the systematic use of positive reinforcement and the elimination of punishment and coercion.

- Train instructional staff in the use of programmed curriculum materials so that each child is enabled to work effectively at his own rate of speed.
- Train parents to work (as paid staff) in the classroom so that they will have the opportunity to influence their children's future education through the use of behavior analysis techniques. (USOE, 1973)

7.0 HIGH/SCOPE EDUCATIONAL RESEARCH FOUNDATION

Cognitively Oriented Curriculum Model

This model represents a synthesis of research in preschool and early elementary education. It focuses on an "open framework" classroom which combines an emphasis on active experience and involvement of the child; a systematic, consistent, and thoroughly planned approach to child development and instruction by the teacher; and continuous assessment of each child's level of development so that appropriate materials and activities can be provided.

The curriculum is cognitively oriented and takes into account the differences between the way children and adults "think." The model's aim is to develop in children the thinking skills they will need throughout their school years and adult lives. The emphasis is on the process of learning rather than a particular subject matter, although the academic subject competencies traditional to the elementary years are taught. The model is an active one. It holds that learning should be active and that it takes place through the child's action on his environment and his resultant discoveries.

Objectives:

- Nurture in the child the thinking and communication skills he will need throughout his school years and his adult life.
- Develop the child's ability to make decisions about what he is going to do and how he is going to do it.
- Develop the child's ability to express himself--to speak, write, dramatize, and graphically represent his experiences and communicate these experiences to others.
- Develop the child's ability to comprehend others' self-expression by reading their writing and understanding artistic and graphic representation.

- Develop the academic subject competencies through application of developing thinking abilities.
- Develop the child's ability to work with other children and adults, so that work done is a result of group planning and cooperative effort.
- Develop the child's self-discipline, his ability to identify personal goals, and to pursue and complete chosen tasks.
- Help the child develop a spirit of inquiry and openness to knowledge and the points of view of others. (USOE, 1973)

8.0 UNIVERSITY OF FLORIDA

Florida Parent Educational Model

This model is based on the premise that it is not enough to change the way the school teaches children; one must also change the way their mothers teach them. Therefore, in this program teaching occurs in both the home and the school and is coordinated by a paid parent educator who comes from the same population as the children's mothers. The parent educator is trained by the project personnel. In the classroom she functions as a teacher's aide, but outside the classroom she instructs mothers in how to teach the child and follow up on his classroom activities. Thus, the mother learns the importance of the home in the child's development and education; she learns what activities to encourage, which to discourage, and perhaps most important, that her actions can have an effect on her child. Mother is encouraged to report to the parent educator which strategies seem to work. She is recruited, therefore, as an active agent in her child's growth and development.

The intrinsic rewards this model yields for the parents are stimulating to the child in that they encourage an environment of pride, achievement, and of high self-esteem. In short, the program seeks to overcome the cycle of despair and low self-concept frequently found in low income populations by encouraging a process of active parent involvement. While the curriculum is not standardized, it does have a Piagetian orientation. The child is encouraged to be experimental rather than repetitious but no particular use of rewards is made. Mastery itself is felt to be its own reward.

Objectives:

- Improve the child's school achievement through work on tasks at home.
- Expand the child's learning environment beyond the school.
- Educate parents to participate directly in the education of their children.
- Motivate parents to develop a home environment that stimulates better performance by the child in school and in life.
- Develop a home-school partnership in all areas of school activities.
- Educate school personnel to support and encourage parental cultural contributions. (USOE, 1973)

9.0 EDUCATIONAL DEVELOPMENT CENTER

EDC Open Education Program

This program is derived from the British Infant School approach, which evolved over the past few decades. It also draws heavily on the knowledge gained in child development over the past 50 years. The approach is essentially a program for helping communities generate the resources to implement open education.

EDC believes that learning is facilitated by a child's active participation in the learning process and that a fundamental educational aim is for children to assume responsibility for their own learning. Learning, therefore, takes place best in a setting where there is a range of materials and problems to investigate which complement the range of ways different children learn.

In an "open" classroom there is a rich environment of materials for children to explore. They are encouraged to initiate activities, be self-directing, and become intensely involved in their interests. Typically there is a variety of activities going on, many of them interdisciplinary. Time is flexible and self-management is the norm, yielding an atmosphere of cooperation where children work together and help each other learn. There may be many interest areas in the room, some reflecting traditional subject matter distinctions such as social studies or mathematics. The classroom is characterized by an interaction of subject matter and purposeful mobility and choice on the part of children.

The role of the teacher in the open classroom is an active one. The teacher leads children to extend their own projects through thoughtful responses and suggestions. This responsive, insightful person enters into the child's growth as a guide who is constantly involved, not as a director or spectator. The objective is to get the children involved in things that are relevant to them. To do this all things are potentially legitimate, although reliance on a structured, prepackaged curriculum is discouraged. The content of what is taught is also rather open, being most strongly influenced by local conditions and objectives. In this approach the emphasis is not so much on content but rather on a process. Within the successful open classroom, learning to take responsibility for one's own learning is perhaps the most important goal.

Objectives:

- Create classroom environments which are stimulating and responsive to a child's individual needs and which make full use of the talents and creative styles of the teachers and aides.
- Develop academic skills in flexible, self-directive ways that allow learning to become part of children's life-styles outside as well as in the classroom.
- Provide resources and environment for children's growth in problem solving skills, ability to express themselves creatively in their social and emotional development, and their ability to take responsibility for their own learning. (USOE, 1973)

10.0 UNIVERSITY OF PITTSBURGH

Individualized Early Learning Program

This program is based on the proposition that if a child is to learn most efficiently he must proceed at his own rate. If a curriculum is to teach most efficiently, components must be carefully, optionally sequenced. The project has, therefore, developed a highly structured and interrelated curriculum. It is based upon a component analysis wherein objectives are stated, requisite skills are specified in behavioral terms, lower level skills are deduced and specified, and higher level skills identified, until a clearly articulated hierarchy has been derived both from logic and a knowledge of psychology. Tests which measure the acquisition of these skills are constructed in a similar manner. These provide a check on the child's progress, the teacher's success, and the adequacy of the component analysis.

There are three general classes of skills included in the curriculum which are felt to underlie all higher order functioning. These are (1) orienting and attending skills; (2) perceptual motor skills, including gross and fine motor skills, such as visual and auditory perception; and (3) conceptual and linguistic skills which include classification, reasoning, memory, language, and early mathematical concepts.

Children learn, the model assumes, by interacting with materials and other children. The teacher serves as a facilitator, monitor, and reinforcer. The teacher, using the diagnostic tests, helps the child move along through the component curriculum using the least powerful reinforcers needed until the child is able to work independently of reinforcement.

Objectives:

- Identify each child's strengths and weaknesses and provide the child with a personal program of instruction based on his individual needs. (USOE, 1973)

11.0 SOUTHWEST EDUCATIONAL DEVELOPMENT LABORATORY

Language Development (Bilingual) Education Approach

This approach is a design for classrooms where a majority of the pupils are Spanish-speaking. The model holds that language is the child's main tool for dealing with his environment, expressing feelings, and acquiring skills, including nonlinguistic ones. An underlying premise is that learning in a second language is easier and more effective if the child first learns concepts and content in his native language and if intensive oral language development in both languages precedes the learning of literary skills. In addition, a positive emphasis on the child's native language and culture is essential to the development of a positive self-concept and pride of heritage.

Step-by-step sequential procedures are followed in teaching language patterns. Both teaching procedures and materials are designed to develop a hierarchy of thinking processes. The focus in teaching language is on content such that all classroom activities reinforce language development.

The model stresses a high level of adult-child contact. Teachers and aides are constant language models giving the child frequent assurance and reinforcement. The kindergarten class, which stresses visual, auditory, and motor skills, as well as thinking, discovery, and English language structures, is divided into small groups which work both independently or with a teacher. In the first and second grades, where oral communication as well as reading and writing skills are stressed, the teacher presents a lesson to the whole group and then the children work independently in small groups.

Objectives:

- Train instructional staff to appreciate the child's culture, to act as good language models, and to become proficient in language development activities.
- Utilize the child's existing concepts as a basis for sequential development of more advanced concepts.
- Teach children to understand, listen, speak, read, and write with equal competence and facility in both the native language and English. (USOE, 1973)

References

Abt Associates Inc. Education as Experimentation: A Planned Variation Model Volume 1A: Early Effects of Follow Through. Contract No. OEC-0-72-5221, the United States Office of Education, Project Follow Through, 1974.

Abt Associates Inc. Interim Report, Volume II. May 1, 1974, Contract No. OEC-072-5221, the United States Office of Education, Project Follow Through, 1974.

Beller, K. B. "Research on Organized Programs of Early Education" in Travers, R.M. (Ed.) Second Handbook of Research on Teaching, Rand McNally & Co., 1973, 530-601.

Cohen, J. & Cohen, P. "Applied Multiple Regression/Correlation Analyses for the Behavioral Sciences", unpublished manuscript, 1973.

Coleman, J. S. The Concept of Equality of Educational Opportunity. Harvard Educational Review, 1968, 38, 7-22.

A Guide to Follow Through. Washington, D.C.: U. S. Office of Education, 1973.

Herriott, R. E. & Hodgkins, B. J. The Environment of Schooling: Formal Education as an Open Social System. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Office of Civil Rights Directory of Public Elementary and Secondary Schools in Selected Districts: Enrollment and Staff by Racial/Ethnic Group. Washington, D.C.: U. S. Government Printing Office, 1970.